United States Court of Appeals for the Federal Circuit

2014-1360

KERANOS, LLC,
Plaintiff-Appellant,
and
UNITED MODULE CORPORATION, PETER COURTURE, and J. NICHOLAS GROSS,
Third Party Defendants,
v.
SILICON STORAGE TECHNOLOGY, INC., FREESCALE SEMICONDUCTOR, INC., MICROCHIP TECHNOLOGY, INC., SAMSUNG SEMICONDUCTOR, INC., SAMSUNG ELECTRONICS CO. LTD., TAIWAN SEMICONDUCTOR MANUFACTURING CO., LTD., and TSMC NORTH AMERICA,
Defendants-Appellees.

Appeal from the United States District Court for the Eastern District of Texas in 2:13-cv-00017-MHS-RSP, Judge Michael H. Schneider

2014-1500

KERANOS, LLC,
Plaintiff-Appellant,
v.
ANALOG DEVICES, INC., APPLE INC., INTERNATIONAL BUSINESS MACHINES CORPORATION, INTEL CORPORATION, NATIONAL SEMICONDUCTOR CORPORATION, NXP SEMICONDUCTORS USA, INC., and TEXAS INSTRUMENTS, INC.,
Defendants-Appellees.

Appeal from the United States District Court for the Eastern District of Texas in 2:13-cv-00018-MHS-RSP, Judge Michael H. Schneider

NON-CONFIDENTIAL OPENING BRIEF OF PLAINTIFF–APPELLANT KERANOS, LLC

James C. Otteson
Michelle G. Breit
AGILITY IP LAW, LLP
149 Commonwealth Drive, Ste. 1033
Menlo Park, CA 94025
650-227-4800
Attorneys for Plaintiff–Appellant

August 11, 2014

CERTIFICATE OF INTEREST

Counsel for the Appellant Keranos, LLC certifies the following:

1. The full name of every party or amicus represented by me is: Keranos, LLC.

2. The name of the real party in interest (if the party named in the caption is not the real party in interest) represented by me is: Not applicable.

3. All parent corporations and any publicly held companies that own 10 percent or more of the stock of the party or amicus curiae represented by me are: None.

4. The names of all law firms and the partners or associates that appeared for the party or amicus now represented by me in the trial court or agency or are expected to appear in this court are:

AGILITY IP LAW, LLP:

James C. Otteson
Michelle G. Breit
Thomas T. Carmack
David A. Caine
John W. Downing
Monica Eno
Xiang Long
Theresa E. Norton

SPANGLER LAW P.C.
(fka Spangler & Fussell, P.C.):

Andrew W. Spangler
James A. Fussell, III

WEISBROD MATTEIS & COPLEY PLLC:

James A. Fussell, II

Dated: August 11, 2014

/s/ Michelle G. Breit
Michelle G. Breit

TABLE OF CONTENTS

Confidential Material Redacted

CONFIDENTIAL MATERIAL OMITTED

Confidential material that is subject to a protective order has been omitted from this public version of the brief. The following information has been redacted:

Brief Page No.	General Description of the Redacted Information
iii	Material omitted describes confidential aspects of defendants, product design and manufacturing methods.
21-24	Material omitted describes confidential aspects of defendants, product design and manufacturing methods. Further, the material omitted includes information derived from defendants' confidential sales information.
25-26	Material omitted describes certain defendants, internal information regarding the source of their products and product naming conventions.
26-27	Material omitted describes certain information about SST's sales of products in the United States.
38	Material omitted describes an aspect of the licensing arrangement between SST and its licensees.
40	Material omitted describes confidential aspects of defendants' product design and manufacturing methods. Further, the material omitted includes information derived from defendants' confidential sales information.
48	Material omitted describes certain information about SST's sales of products in the United States.
49	Material omitted describes certain information about SST's sales of products in the United States.
51	Material omitted describes confidential aspects of defendants' product design and manufacturing methods.

TABLE OF AUTHORITIES

Cases

Statutes

Rules

STATEMENT OF RELATED CASES

This Court's Order of July 11, 2014 consolidated the two appeals listed in the caption. There are no other cases in this or any other court known to directly affect or be directly affected by this Court's decision in the instant consolidated appeal. There have been no appeals in either of the consolidated cases that were previously before this or any other appellate court.

JURISDICTIONAL STATEMENT

The district court had jurisdiction over Keranos's patent infringement actions under 28 U.S.C. §§ 1331 and § 1338. This Court has jurisdiction under 28 U.S.C. § 1295(a)(1). These appeals are from final judgments under Rule 54(b) entered on February 10, 2014 and May 1, 2014. A1; A17. The appeals in the consolidated cases were timely filed on March 7, 2014 and May 9, 2014. *See*, A79; A116-117.

I. STATEMENT OF THE ISSUES

A. Whether the district court abused its discretion in finding that Keranos's infringement contentions failed to comply with Eastern District of Texas Patent Local Rule 3-1 where: (i) the infringement contentions identified the accused products by the trade name defendants use for the infringing memory cells ("SuperFlash®"), the name of the three generations of cell design architectures ("ESF-1," "ESF-2," and "ESF-3"), as well as representative product or family

numbers known to Keranos at the time; (ii) defendants and their expert agreed that all SuperFlash memory cells (within a given generation) were manufactured and functioned in exactly the same way; and (iii) defendants admitted that it was impossible for Keranos to use publicly available sources to identify all specific product names/numbers for all products with SuperFlash memory cells.

B. Whether the district court abused its discretion in finding that Keranos was not diligent in identifying accused products where: (i) Samsung did not oppose Keranos's motion to amend its infringement contentions, and there is no evidence that Keranos could have identified Samsung's product numbers from public sources; (ii) Keranos's infringement contentions for TSMC used TSMC's trade name ("EMBFLASH") for the accused SuperFlash memory cells, and TSMC admitted that Keranos could not identify TSMC's product numbers through public sources; and (iii) the other defendants admitted that publicly available information only identified their products generically as having "flash" memory, but did not indicate which products included "SuperFlash" memory cells.

C. Whether the district court abused its discretion in finding that Keranos lacked good cause to amend its infringement contentions where (i) the excluded products were of great importance in that they constituted nearly all the accused products, and Keranos was time-barred from asserting claims against those products in a new action; (ii) defendants would have suffered no prejudice if the

amendment were allowed; and (iii) no continuance would have been necessary had the district court granted the motion.

D. Whether the district court abused its discretion in effectively dismissing Keranos's patent infringement claims as a sanction for non-compliance with Patent Local Rule 3-1 where Keranos did not act willfully or in bad faith, defendants suffered no prejudice, and lesser alternative remedies were available without resort to drastic sanctions effectively depriving Keranos of its patent infringement claims against defendants.

II. STATEMENT OF THE CASE WITH RELEVANT FACTS

A. Introduction.

Keranos asserts that defendants infringed the patents-in-suit (directly and/or indirectly) by making, using, selling and/or importing wafers or integrated circuits ("ICs") (or products containing ICs) that include flash memory cells designed by defendant SST called "**SuperFlash**®." In its Patent Local Rule 3-1 ("Rule 3-1") infringement contentions ("IFCs"), Keranos identified the accused ICs by the name "SuperFlash," which defendants themselves use to identify the infringing memory cells. The IFCs also included the commercial names defendants use to identify the design architecture for the accused SuperFlash memory cells, and the product numbers and/or family names/numbers known to Keranos. Defendants did not move to strike (or seek other relief) regarding Keranos's IFCs at any time after receiving service of the IFCs.

Through its first discovery requests, Keranos obtained the identity of additional part/family numbers under which defendants sold products with SuperFlash memory cells. After that, Keranos promptly moved to amend its IFCs to reflect those product numbers. Defendants opposed Keranos's motion claiming Keranos lacked diligence, while never asserting Keranos's IFCs failed to provide them with sufficient notice of Keranos's infringement theories.

One year later, after the parties completed extensive discovery regarding all products with SuperFlash cells that Keranos listed in its proposed amended IFCs, and on the eve of trial, the district court denied Keranos's motion to amend its IFCs. The court ruled that:[1] (i) Keranos's method of identifying the accused products in its IFCs did not comply with Rule 3-1; (ii) Keranos could not amend the IFCs because it lacked diligence by failing to identify defendants' individual products by their separate product numbers from public sources and without discovery; and (iii) the only products deemed in the litigation were those the IFCs identified by individual product number. A13; A7-8; A506-8 (6:13-8:20). The district court's order precluded Keranos from asserting its claims against all but a *de minimis* portion of the thousands of accused products. Because the patents-in-suit had expired and damages were time-barred, the district court's

[1] The court's rulings on the motion to amend are stated among the court's order denying Keranos's motion to amend, order denying Keranos's motion for reconsideration, and clarification of the orders Magistrate Judge Payne made from the bench during a status conference held January 29, 2014.

belated order amounts to a "death sentence," prohibiting Keranos from ever pursuing its claims.

As set forth below, the district court abused its discretion in finding that Keranos's original IFCs did not comply with Rule 3-1 or alternatively in denying leave for Keranos to file its amended IFCs. Keranos's IFCs provided defendants with more than ample notice and information regarding the identity of the accused instrumentalities and its infringement theories with which to litigate their defenses – the very purpose of Rule 3-1. Keranos's proposed amended contentions did not alter the scope of the products Keranos accused, nor did they change Keranos's original infringement theories in any way. Even if the IFCs did not comply fully with the letter of Rule 3-1, the court's conclusion that Keranos lacked good cause to amend was based on findings unsupported by the record, including erroneous speculative findings that Keranos could have identified thousands of product numbers through public sources and without discovery.

Beyond that, the draconian sanction the court imposed for Keranos's purported failure to comply with the district court's scheduling order or local discovery procedure – *i.e.*, dismissal of essentially all of Keranos's infringement claims – is not only unjust, it is inconsistent with the Federal Rules of Civil

Procedure, which prohibit such drastic sanctions in the absence of bad faith or prejudice or where less drastic sanctions are available.

B. Overview of Case Histories and Alignment of Parties.

Keranos initiated a patent infringement lawsuit on June 23, 2010 in the Eastern District of Texas against various parties including each of the defendants in the consolidated appeals herein, with the exception of SST (the "Original Case").[2] *See*, A22879-22992. In late 2010, SST and certain defendants in the Original Case filed complaints in the Northern District of California [*see*, A22993-23068; A23069-23150; A23151-23230; A23231-23307], seeking declarations of noninfringement and invalidity of the same patents in the Original Case. The declaratory judgment actions were transferred to the Eastern District of Texas, where the court eventually consolidated them with each other (the "DJ Case").[3] S*ee*, A198 (Dkt. 85).

Following a joint *Markman* hearing in the Original and DJ Cases on December 12, 2012, the district court regrouped the accused infringers into two new cases: (1) one with the manufacturers and sellers of semiconductors with

---

[2] *Keranos LLC v. Analog Devices, Inc., et al*., E.D. Tex. Case No. 2:10-cv-00207.

[3] *Samsung Electronics Co., LTD, et al. v. Keranos LLC*, E.D. Tex. Case No. 2:11-cv-00331 (consolidated with Case Nos. 2:11-cv-00332, 2:11-cv-00333, and 2:11-cv-00334).

SuperFlash memory cells (the "Manufacturer Case")[4]; and (2) the other with customers of SuperFlash semiconductor products many of which incorporated those products into larger products for sale (the "Customer Case").[5] Thereafter, the district court treated the Manufacturer Case as the lead case. Discovery continued in the Manufacturer Case, but was deferred in the Customer Case.

C. The Patents-In-Suit.

Keranos is the exclusive licensee of all substantial rights to the three patents-in-suit: U.S. Patent Nos. 4,868,629 ("the '629 patent") [A259-286]; 4,797,719 ("the '719 patent") [A247-258]; and 5,042,009 ("the '009 patent") [A287-304], which originated from work done on flash technologies by pioneering inventors at Waferscale Integration. The three patents relate to memory cells in semiconductor chips, and methods for manufacturing and programing those memory cells. In particular, the patented memory cells include split-gate transistors, which allow a

---

[4] Case No. 2:13-cv-00017 includes defendants Freescale Semiconductor, Inc. ("Freescale"), Microchip Technology, Inc. ("Microchip"), Samsung Semiconductor, Inc. and Samsung Electronics Co., Ltd (jointly "Samsung"), SST, and Taiwan Semiconductor Manufacturing Company, Ltd. and TSMC North America (jointly "TSMC") (collectively the "Manufacturer Defendants").

[5] Case No. 2:13-cv-00018 includes Analog Devices, Inc. ("ADI"), Apple, Inc. ("Apple"), International Business Machines Corporation ("IBM"), Intel Corporation ("Intel"), National Semiconductor Corporation ("NSC"), NXP Semiconductors USA, Inc. ("NXP") and Texas Instruments, Inc. ("TI") (collectively the "Customer Defendants").

"charge" on the transistor to be stored, read and erased. A15654-5. They retain their programmed state even when the electronic device containing the transistors is powered off. A15655. This is especially advantageous because the split-gate memory cells can be embedded in larger ICs such as microcontrollers, since they are compatible with the processing required by such devices.

The '719 and '629 patents solved a significant limitation in the prior art: how to make a split-gate cell with a predictable channel length underlying the floating gate. A268, col. 3:57-61. Controlling the channel length, particularly the portion under the floating gate, is critical for an optimal operation of the split gate cell. A267-8, col. 2:57-3:21. The patents describe a superior way of manufacturing split gate cells where one edge of the floating gate is used as a mask to align the drain region. The innovative inventions significantly improved yield caused by misalignment manufacturing errors in prior art memory cells. A269, col. 6:9-20. The resulting memory cell is a more efficient, predictable, and reliable split-gate transistor. The '719 patent claims new methods of manufacture of split-gate memory cells and the '629 patent claims the split-gate memory cells themselves.

The '009 patent teaches methods for efficiently programming split-gate non-volatile memory cells. To program a non-volatile memory cell, a charge must be placed on the floating gate. The floating gate is insulated from the rest of the

transistor and therefore can retain or store a charge even when the voltages applied to the control gate and drain region are turned off. Because memory cell arrays consist of millions of individual cells, information can be readily stored and retrieved on electronic devices.

The methods claimed in the ’009 patent relate to hot-electron injection programming. In general, hot electron injection programming consists of applying voltages to the control gate and drain regions of the memory cell. The voltages create a current flowing in the channel region between the source and the drain regions. If the voltage and programming drain current are sufficiently large [A298, col. 1:50-55], some of the electrons in the current will reach an excited state and inject themselves onto the floating gate, placing a charge on it.

The ’009 patent inventors recognized that at the time of the invention, hot electron injection programming required high voltages and high programming current to create the conditions necessary for hot electron injection. A298, col. 1:26-2:2. The inventors discovered and claimed methods “for programming without providing such a large drain current.” A298, col. 5:14-16. Thus, the ’009 patent teaches several ways to program memory cells so that an on-chip power supply with limited power can supply the necessary current. A267, col. 2:24-31.

D. SuperFlash Technology and the Defendants.

"SuperFlash" is a flash memory technology developed by SST based on split-gate transistors, and marketed under that tradename. According to SST, SuperFlash® split-gate memory is desirable because of its simplified design, low power, high performance and scalability. A6321; A6334. Over the course of several years, SST developed three basic SuperFlash® memory cell structures, which SST refers and licenses to the Manufacturing Defendants under the names "ESF1" (first generation), "ESF2" (second generation) and "ESF3" (third generation). A1131, Fig. 2. Every product accused of infringement in this action uses one of these three basic structures.

SST itself sells IC products that include embedded SuperFlash memory. SST also licenses its SuperFlash technology to IC manufacturers. A5885-5890. Under its licenses, SST transfers specific SuperFlash know-how to its licensees to facilitate their installation and use of the technology. A6499 (232:14-18); A5885-5890; *see also*, A6223-4, A6226-7. The Manufacturer Defendants are all SST licensees, each of which has made, sold or imported wafers (TSMC) or semiconductor chips (the remaining Manufacturer Defendants) with SuperFlash memory. *Id.*; *see also*, A22861. The Customer Defendants have each purchased SuperFlash memory devices and incorporated them into their products which they sold or imported into the United States. A22861-2; A6231-3.

E. Keranos's Pre-Filing Investigation.

Before filing its complaint (and serving its IFCs), Keranos spent hundreds of hours to identify products that utilize SuperFlash® memory cells (or derived from SuperFlash technology). A1142. Keranos reviewed defendants' websites (current and archived), performed parametric searches, and reviewed defendants' datasheets. A1142. Keranos also reviewed SST's annual reports, presentations and press releases to identify companies whose products included SST SuperFlash memory cells. A1142; *see also*, *e.g.*, A5885-5890.

The process of identifying individual infringing products was not straightforward. Although publicly available information identified companies whose products included SuperFlash technology, for most defendants, publicly available material did not explicitly identify which of their numerous products implemented SuperFlash memory cells. A1142. Instead, much of the publicly available information Keranos found through its investigation merely noted generically that the IC products had "flash" or "embedded flash" memory. A1142.

Keranos encountered other obstacles in identifying product names/numbers, because the recoverable damages period included only products from 2004 through 2008. In the limited instances where a defendant's materials identified a particular product as SuperFlash, Keranos could not reliably confirm the presence of SuperFlash during the relevant timeframe. A1143. For example, absent discovery,

Keranos found no practical mechanism to verify that a product advertised for sale in 2010 was also for sale between 2004 and 2008 ***and*** employed SuperFlash technology during that time. A1143.

During their hundreds of hours of investigation, Keranos and its experts reviewed hundreds of datasheets. A1143. As a result of its efforts, Keranos was able to identify some products by part number or product family name that defendants advertised as including SuperFlash memory. A1142. As to products that defendants did not identify publicly as including SuperFlash, Keranos made educated guesses regarding a sampling of products. A1143. From there, Keranos purchased commercial reverse-engineering product reports and hired experts to reverse-engineer the products (confirmed to have been manufactured in the relevant time frame), which Keranos's experts believed were good candidates for utilizing SuperFlash. A1143. Keranos approximates that it spent well over $30,000 in fees for reverse-engineering products. A1143. Keranos's reverse-engineering efforts were not always successful. Keranos spent thousands of dollars reverse-engineering, only to find in instances that certain products did not utilize the split-gate SuperFlash cells at issue in this case. A1143.

F. Initial Phase of the Litigations.

Keranos filed its original complaint on June 23, 2010. A22879-22992. Thereafter, the litigation stalled while the parties engaged in substantial procedural

disputes and motion practice, mostly initiated by defendants. In various motions, defendants asserted issues regarding: lack of standing; failure to state a cause of action; failure to adequately plead indirect and willful infringement; and lack of personal jurisdiction, among other assertions. *See*, A151-171. During the same time, Keranos asked the Northern District of California – where SST and some of the Original Case defendants had improperly filed declaratory judgment actions – to transfer those cases to Texas. The DJ Cases were transferred to the Eastern District of Texas around July 21, 2011. A202 (Dkt. 40); A215 (Dkt. 85); A231 (Dkt 35); A242 (Dkt. 40). On September 11, 2011, the district court ruled on the motions in the Original Case, denying most of defendants' requests. A23308-23328.

By district court order, the parties in the Original Case then met and conferred and filed a proposed docket control order and proposed discovery order on October 21, 2011. A147 (Dkt. 442). The docket control order entered on October 26, 2011, required Keranos to serve its IFCs on or before December 16, 2011, which date the court later extended to December 19, 2011. A23329-23333; A23334. The parties agreed that Keranos could begin discovery in January 2012, after it served its IFCs in December 2011. In the DJ Cases, pursuant to a court order, the parties submitted a proposed scheduling order on March 2, 2012, which the court entered on March 13, 2012. A23335-23338.

G. Keranos's Rule 3-1 Infringement Contentions.

In accordance with the district court's order in the Original Case, on December 19, 2011, Keranos timely served its IFCs. The IFCs consolidated in one document its contentions as to all accused infringers in the pending cases. A21541-631. Keranos also served its IFCs in the DJ Cases, although the court had not yet scheduled a date for service of IFCs in those cases.

Discovery had not yet opened and Keranos relied on the information known to it after conducting hundreds of hours of research and obtaining a number of reverse-engineering reports. The IFCs identified each defendant's accused products by reference to "SuperFlash" (the trade name defendants used for the accused technology) and the names of the specific SuperFlash architectural designs defendants used in their products, *i.e.*, "ESF1," "ESF2" and "ESF3." The IFCs defined the accused "SuperFlash® ICs" (integrated circuits) as:

> ICs [integrated circuits] embodied in discrete form or wafer form and having an infringing flash memory portion (including standalone memory and embedded memory types referred to as "SuperFlash" memory) originating from an SST design (including any one of at least three separate generations of cells identified as ESF1, ESF2 and ESF3 by SST) or a derivative thereof ("SuperFlash® ICs").

A21544.

In other words, Keranos tailored the IFCs to specifically identify the accused products as defendants' (i) integrated circuits (or products containing the same), (ii) in discrete or wafer form, (iii) having SuperFlash memory, and (iv) using

SST's first, second or third generation memory cell structures identified by name as ESF1, ESF2 and ESF3, respectively, and products that incorporated the same. In addition, for all defendants, the IFCs identified one or more specific part numbers or product family names under which each defendant sold SuperFlash products. Further, pursuant to Rule 3-1(c), the IFCs included highly detailed claim charts showing the structure of each of the accused SuperFlash ESF1, ESF2 and ESF3 memory cell designs, and describing exactly where in each memory cell the claim limitations of the patents-in-suit were found and how the accused SuperFlash® products practiced the claimed methods. A21567-85; A21587-611; A21613-31. The claim charts included annotated figures and drawings and electron micrographs of reverse-engineered products. Because the structure and operation is exactly the same, Keranos presented the same infringement theory against each of defendants' SuperFlash products identified in the IFCs.

Notably, after receiving the IFCs, not a single defendant moved to strike them for failing to comply with Rule 3-1 or otherwise sought relief from the court based on assertions that the IFCs were deficient in any way. In particular, no defendant argued that Keranos's infringement theories were insufficient for defendants to fully and fairly defend against Keranos's claims.

H. Keranos's Motion to Amend its Infringement Contentions.

1. Keranos's Efforts Leading to the Motion to Amend.

Discovery opened in the Original Case in January 2012, and Keranos promptly served written discovery on the defendants in that case, including interrogatories seeking information regarding all products using SuperFlash® memory. A1070; A22554-61. In their initial responses, most of the defendants failed or refused to provide any discovery except as to the SuperFlash ICs that Keranos specifically identified by family number or part number in its IFCs. *See*, A2013-22 (ADI); A22457-71 (Apple); A15520-33 (IBM); A2256-2266 (Intel); A2048-57 (NSC); A2268-84 (NXP); A2059-68 (TI); A2036-46 (Freescale); A1945-2011 (Microchip); A2070-80 (Samsung); A6216-42 (SST); A2024-34 (TSMC). Microchip, TSMC, and ADI were the exceptions who, in March 2012, provided discovery responses that identified their SuperFlash® products by product number. A1945-2011 (Microchip); A2024-34 (TSMC); A2013-22 (ADI). The other defendants continued to assert objections and Keranos undertook meet and confer efforts with those parties.

In April 2012, Keranos provided the Original Case defendants with notice that it intended to amend its IFCs to include product numbers obtained through discovery. A6103. A few days later, Keranos circulated a proposed amended Rule 3-1 pleading that reflected the product numbers that the few defendants had

disclosed by that time. A2091-221. Further, after the court entered the scheduling order in the DJ Cases, Keranos served its first round of discovery on SST. A1028-1030.

Keranos continued to meet and confer with the remaining defendants so that it could amend its consolidated IFCs as a single amended document. *See*, *e.g.* A2246-54; A22135-37; A2082-85; A22139-40; A1459-61; A2223-25; A2227-28; A6093-95; A22146; A22563-66; A22148-52; A22154-60; A2230-333; A22162-64; A2235-36; A6097-98; A22166; A22168-9; A4219-20; A22171; A22173-74. Apple provided a list of product numbers on May 17, 2012 [A22148-52] but did not supplement its discovery responses to reflect the numbers until June 28, 2012. NSC, TI and NXP did not disclose product numbers or product family numbers until June 15, 2012. *See, e.g.*, A3720-26; A5892-910. SST disclosed the products – by product *family* numbers only – on June 18, 2012. A6221-24. On June 20, 2012, Samsung disclosed its SuperFlash product numbers [A1815], and Intel disclosed product numbers on June 22, 2012. A18791-92. Freescale did not disclose its product numbers until July 2, 2012. A5429-75. IBM never identified IBM products that included SuperFlash® ICs beyond those Keranos specifically identified in its IFCs; on July 20, 2012, IBM instead identified several SST SuperFlash® products it had purchased, without revealing in which IBM products it used the ICs. A15520-33.

Immediately after obtaining information from all defendants, on July 20, 2012, Keranos moved the district court for leave to amend its IFCs to reflect the product numbers it learned through discovery. A1121-1462; A1463-1805. At that time, the parties had engaged in minimal written discovery; claim construction briefing was months away; the parties had not noticed (much less conducted) any depositions; no experts had been disclosed or expert opinions rendered; and the court had set no trial dates.

2. Defendants' Responses to Keranos's Motion to Amend.

In the Original Case, Freescale, Microchip, TSMC, ADI, NSC, NXP, and TI (among other parties no longer in the litigation) filed a joint response in opposition to Keranos's motion to amend. A1806-3790. In the consolidated DJ Case, SST, Freescale, ADI, Microchip, TSMC and NSC filed a joint response in opposition. A3791-5771; *see also*, A18967-20946. Intel, IBM and Apple each filed separate opposition memoranda. A18775-18783 (Intel); A18949-18966 (IBM); A20953-21645 (Apple). Samsung filed no opposition at all. Even in opposing the motion to amend, none of the defendants contended that the IFCs provided insufficient notice of the accused products or that the IFCs failed to adequately disclose Keranos's infringement theories.

Rather than address the adequacy of the notice and information the IFCs provided to them, most defendants argued Keranos had not been diligent because it

had not discovered defendants' individual product numbers on its own before serving the IFCs. Strictly crediting form over substance, those defendants pointed to Rule 3-1(b)'s language that product "identification shall be as specific as possible" and the accused products should be "identified by name or model number, if known." They argued Keranos could have discovered, on its own by searching public sources or conducting additional reverse-engineering, the individual product numbers defendants assigned to the thousands of products using SuperFlash® technology. Defendants argued the purported deficiency prohibited Keranos from amending its IFCs and thereby from pursuing its claims against them.

In fact, only SST, Freescale and Apple pointed to publicly available information that both identified specific products by individual product number ***and*** confirmed that the products include SuperFlash cells. Freescale, however, acknowledged that such public information existed for only for a fraction of its SuperFlash® products. A1823. As for the majority of the Freescale products that Keranos sought to list in its amended IFCs, Freescale further admitted that public information, such as datasheets, indicated only that the products generically included "flash memory." A1823.

Microchip, ADI, NSC, TI and NXP, IBM, Intel could not point to any publicly available information identifying their products by product number and

stating the specific products contain "SuperFlash" memory. Like Freescale, defendants Microchip, ADI, NSC and TI pointed only to public disclosures that their products generically contained "flash memory." A1824-28. None of these defendants provided an explanation as to how Keranos could have identified which ones of the multitude of "flash memory" products included SuperFlash® memory, as opposed to other flash memory. Instead, they speculated that because Keranos concluded some such products included SuperFlash®, Keranos should have alleged that all of defendants' "flash memory" products included SuperFlash® memory and identified all those products in the IFCs by their part numbers. *Id.*

TSMC admitted that Keranos could not identify TSMC's product numbers through public sources. A1819. Instead, TSMC argued Keranos should have moved sooner to amend to add TSMC's product numbers disclosed in discovery.

I. Further Conduct of the Case, Including Relevant Discovery and Expert Reports.

1. The Parties' Continued the Litigation as if All SuperFlash® Products Were Part of the Case.

The district court did not rule on Keranos's motion to amend until more than a year after Keranos filed it, and just weeks before the pre-trial conference. Tellingly, after Keranos filed the motion, all parties conducted the litigation as if all SuperFlash products – including all products identified in Keranos's proposed amended IFCs – were part of the case. During that time, the parties in the

Confidential Material Redacted

Manufacturer Case completed discovery as to all the SuperFlash products. They conducted 30 depositions, including in Taiwan and Korea, and various locations in the United States. A6564. The Manufacturer Defendants produced lengthy product and sale summaries regarding all the SuperFlash products. *See, e.g.*, A6923-8054. The parties in the Manufacturer Case exchanged opening and rebuttal technical and damages expert reports and took expert depositions addressing all SuperFlash products. After filing its motion to amend and before the district court ruled, Keranos incurred hundreds of thousands of dollars in expert and support costs alone prosecuting its claims against all SuperFlash products identified in the proposed amended IFCs.

2. [redacted]

Confidential Material Redacted

Confidential Material Redacted

[redacted]

6 [redacted]

Confidential Material Redacted

Confidential Material Redacted

3. Discovery Confirmed that Publicly Available Information Was Not Sufficient for Keranos to Identify Individual Product Numbers for Most SuperFlash® Products.

Discovery responses provided by the Manufacturer Defendants confirmed that the number of SuperFlash® ICs sold having separate part names/identifiers was enormous (thousands), and the part numbers themselves did little (if anything) to reveal to an outsider whether or not the products included SuperFlash® memory. For example, Microchip produced a list of [redacted] separate part numbers for its SuperFlash® ICs. A6023-69. [redacted]

Confidential Material Redacted

[redacted] TSMC does not sell branded ICs with markings, but rather only bulk wafers; thus it would have been impossible for Keranos to "identify" such TSMC products more specifically by number.

Other discovery confirmed Keranos could not know from public sources which products containing SuperFlash® memory were ever made, sold or imported into the United States. [redacted]

Confidential Material Redacted

J. The District Court's Denial of Keranos's Motion to Amend.

On August 5, 2013, more than one year after Keranos filed its motion, the district court entered its order in the Manufacturing Case, denying the motion to amend.[7] A5-8. In the order, the district court found that good cause to amend did not exist because Keranos was not diligent in identifying the individual SuperFlash® product numbers in its original IFCs. A8. According to the court,

[7] The district court wrote in a footnote that following the claim construction hearing, held in December 2012, the court denied the motion to amend without prejudice to refile. Keranos is unaware of such an order and, as it does not appear anywhere in the court's record, submits that this was an oversight.

Keranos could have and should have identified the individual product names from public information in its possession or on the Internet. *Id.* The court construed Keranos's proposed amendment as "add[ing] thousands of additional products not specifically disclosed in Plaintiff's original infringement contentions." *Id.* The district court did not separately address and made no distinction regarding the arguments and evidence (or the lack thereof) relating to the circumstances and facts applicable to the separate individual defendants. The district court simply denied Keranos's motion to amend in its entirety, and applied its order uniformly to all Manufacturing Defendants.

Keranos filed a motion for reconsideration, pointing out new evidence disclosed during discovery which demonstrated the identicality of all SuperFlash® products within a design generation. A6367. Keranos also sought clarification as to: (1) whether the order excluded Keranos from proceeding against products in specific product families that Keranos did in fact identify in its IFCs by family number or family name; and (2) whether the order applied to parties who did not meet this criteria – *i.e.*, Samsung (who did not oppose the motion to amend) and TSMC (who admitted its product numbers were not publicly available). A6566-7.

The district court denied Keranos's motion for reconsideration and categorically ruled that Keranos could not proceed against products within named product families because, according to the court, disclosure of product families

without "attempt to narrow the accused product descriptions is inconsistent with the rules of this court." A13. The district court did not address the record regarding the applicability of this ruling to parties such as Samsung or TSMC.

Notwithstanding the court's denials of the motion to amend and for reconsideration, an ambiguity remained regarding the status of Keranos's infringement contentions. Although the Court had denied Keranos's motion to amend, the Court had not stricken Keranos's infringement contentions; made any specific findings as to which products, if any, Keranos could pursue at trial; or dismissed or otherwise dispositively ruled on the viability of Keranos' patent infringement claims.

The court-appointed Magistrate held a hearing on January 29, 2014, to address these issues among other status-related issues. At the hearing, Magistrate Judge Payne specifically clarified that the only products deemed still at issue were those that Keranos identified by individual product number. A506-8 (6:13-8:20). The district court's order ultimately resulted in only two products remaining in the Manufacturing Case, one SST product, which Keranos had identified in its IFCs but SST later disclosed was never sold in the U.S., and one Microchip product for which damages from the low volume of sales were *de minimis*, *i.e.*, no more than $800. A15625; A508. Keranos dismissed with prejudice its claims against the two products and, based on the court's order, stipulated to summary judgment in favor

of the Manufacturing Defendants. A532-533 (32:23-33:7); A15644-7. The district court then granted summary judgment and entered final judgment against Keranos. A14; A15; A1.

On January 28, 2014, the district court entered an order denying Keranos's motion to amend in the Customer Case, referencing its order in the Manufacturer Case. A18. The parties disputed whether certain products remained in the case and submitted their dispute and arguments to the district court. A22592-22651 (Keranos); A22818-22854 (IBM); A22687-22748 (Apple); A22749-22817 (TI). The court ruled that Keranos's identifications were to product families, and thus not sufficiently definite under the court's orders. A19-24.

As with the Manufacturer Case, the court's application of its order in the Customer Case prohibited Keranos from asserting its patent infringement claims against all but a *de minimis* number of the customer's SuperFlash products. Keranos again agreed to dismiss the *de minimis* remaining products and, based on the court's orders denying Keranos's motion to amend, Keranos stipulated to summary judgment in favor of the Customer Defendants. *See*, A22853-7. The district court then granted summary judgment and entered final judgment against Keranos. A25-27; A17.

III. SUMMARY OF ARGUMENT

The district court abused its discretion in ruling that Keranos's IFCs do not comply with Rule 3-1 and Keranos could not amend the IFCs to reflect the individual product numbers. The district court's conclusion that Keranos's IFCs did not comply with Rule 3-1 misconstrues the rule. Infringement contentions under Rule 3-1 must be sufficient to provide notice of plaintiff's infringement claims. Keranos's identification of the accused products more than met the notice function of Rule 3-1 and identified the products by the commercial name which defendants use to identify the accused instrumentalities.

Even if, as the court concluded, Keranos (i) lacked diligence in not identifying defendants' separate product numbers and, therefore, (ii) did not comply fully with the letter of Rule 3-1, the district court abused its discretion in finding that good cause did not exist for Keranos to amend its IFCs. The district court failed to consider the remaining highly relevant factors used to determine whether good cause exists, including: (a) the importance of the products excluded; (b) the lack of prejudice to defendants; and (c) the absence of need for a continuance – each of which weighs exclusively in Keranos's favor.

In addition, the district court erroneously applied its ruling indiscriminately to all defendants and all accused products, even though the record is devoid of evidence that Keranos could have identified product numbers from public sources

as to all such defendants. Most notably, Samsung did not oppose Keranos's motion to amend, and provided no evidence that Keranos could have identified specific Samsung SuperFlash® product numbers from public sources. TSMC admitted that Keranos could not have learned TSMC's internal product numbers through public sources, and further admitted that the product family identifier Keranos used in its IFCs, *i.e.*, "EMBFLASH," is the exact and precise product name TSMC uses to publicly advertise its SuperFlash® products. Reversal is assuredly required as to those defendants and products for which the record contains no evidence upon which the court rationally could have based its decision.

Similarly, the court abused its discretion in denying Keranos's motion to amend to identify SuperFlash products for which even the only "publicly" available information indicated the products included "flash memory," but with no specific mention of "SuperFlash." By admission of defendants, those products included a majority of Freescale's products and all of Microchip's, ADI's, NSC's, and TI's products. Under the circumstances, the record cannot support a conclusion that Keranos lacked diligence in not identifying these product numbers in its IFCs.

Finally, the district court's interpretation and enforcement of Rule 3-1, which imposed the drastic sanction of effective dismissal on Keranos for its purported failure to comply with the local patent rule, flatly conflicts with the

Federal Rules of Civil Procedure, which prohibit imposition of this "remedy of last resort." Keranos engaged in no willful misconduct or bad faith, timely provided highly detailed IFCs believing they complied with Rule 3-1, and defendants suffered no prejudice by Keranos's purported noncompliance. Indeed, discovery had already been completed as to all points on ALL of the products identified in the amended IFCs when the Court issued its ruling. The court's dismissal of Keranos's claims, without resort to a lesser remedy that included an opportunity to amend its IFCs, was unjust and an abuse of discretion.

For each of these reasons, as set forth in further detail below, Keranos respectfully requests this Court reverse the district court's order.

IV. ARGUMENT

A. Applicable Standards of Review.

The Federal Circuit applies its law in reviewing questions pertaining to the validity and enforcement of patent local rules. *O2 Micro International Limited v. Monolithic Power Systems, Inc.*, 467 F.3d 1355, 1364 (Fed. Cir. 2006) (holding because "local patent rules on amendment of infringement contentions are unique to patent cases and have a close relationship to enforcement of substantive patent law, we proceed to review the validity and interpretation under Federal Circuit law."). The Federal Circuit will reverse a decision enforcing local rules in patent cases if the Court finds the district court abused its discretion. *Id*. at 1366-67. The

district court abuses its discretion when its finding is "clearly unreasonable, arbitrary, or fanciful; based on erroneous conclusions of law; clearly erroneous; or unsupported by any evidence." *Id.* "A finding is 'clearly erroneous' when – although there is evidence to support it – the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed." *United States v. United States Gypsum Co.*, 333 U.S. 364, 395 (1948); *1st Media, LLC v. Electronic Arts, Inc.*, 694 F.3d 1367, 1372 (Fed. Cir. 2012).

In reviewing discovery sanctions or sanctions for non-compliance with a scheduling order, the Federal Circuit applies the law of the regional circuit to which district court appeals normally lie, unless the issue pertains to or is unique to patent law. *Tennant Co. v. Hako Minuteman, Inc.*, 878 F.2d. 1413, 1416 (Fed. Cir. 1989) (vacating dismissal sanctions imposed for discovery abuse); *see also*, *University of Pittsburg v. Varian Medical Systems, Inc.*, 569 F.3d 1328, 1334 (Fed. Cir. 2009) (vacating dismissal sanctions imposed for violation of scheduling order). A decision to sanction a litigant for discovery violations or failure to follow a court's scheduling order is one that is not unique to patent law, and the Federal Circuit therefore applies the regional law as to that issue. *Tennant,* 878 F.2d at 1416; *University of Pittsburg*, 569 F.3d at 1334. In the Fifth Circuit, the district court's imposition of sanctions is reviewed for abuse of discretion. *Tollett v. City of Kemah*, 285 F.3d 357, 363 (5th Cir. 2002). A district court abuses its

discretion when the "ruling is based on an erroneous view of the law or on a clearly erroneous assessment of the evidence." *Id*. (internal quotation marks and citation omitted).

As shown below, the district court's rulings - finding Keranos's IFCs deficient and denying leave to amend - results from applying incorrect legal standards, clearly erroneous fact findings, and arbitrary decision making. In addition, the district court abused its discretion by imposing the most drastic sanction possible – terminating Keranos's claims.

B. Keranos's Infringement Contentions Complied with Rule 3-1.

1. Rule 3-1 is Intended to Provide Parties with Adequate Notice and Information with which to Litigate their Cases.

"The overriding principle of the Patent Local Rules is that they are designed [to] make the parties more efficient, to streamline the litigation process, and to articulate with specificity the claims and theory of a plaintiff's infringement claims." *Bender v. Maxim Integrated Products, Inc.*, No. C-09-01152, 2010 WL 1135762 (N.D. Cal. Mar. 22, 2010) (ordering amend infringement contentions). Stated another way, "[t]he Patent Local Rules 'exist to further the goal of full, timely discovery and provide all parties with adequate notice and information with which to litigate their cases.'" *Nidec Corp. v. LG Innoteck Co., Ltd*, No. 6:07-cv-108, 2009 WL 3673253, at *1 (E.D. Tex. Sept. 2, 2009), *quoting Computer Acceleration Corp. v. Microsoft Corp.*, 503 F.Supp.2d 819, 822 (E.D. Tex. 2007)

(internal citations omitted). "These contentions must be specific enough to give a defendant notice of plaintiff's infringement claims and must go beyond that provided by the mere language of the patent." *Id*. "Infringement contentions are not meant to require a party to prove its case of infringement or provide a forum for litigation of the substantive issues, however." *H-W Technology, L.C. v. Apple, Inc.*, No. 3:11-cv-651, 2012 WL 3650597, *2 (N.D. Tex. Aug. 2, 2012) (citations omitted). "[T]hey are merely designed to streamline the discovery process." *Id*. (citations omitted).

Rule 3-1(b) requires a patentee to identify "each accused apparatus, product, device, process, method, act, or other instrumentality ('Accused Instrumentality') of each opposing party **of which the party is aware**." Rule 3-1(b) (emphasis added). Although the rule provides the "identification shall be as specific as possible," it goes on to explain that each Accused Instrumentality must be "identified by name or model number, **if known**." *Id*. (emphasis added). The rule expressly refers to the **patentee's** awareness and knowledge of each Accused Instrumentality and its name or model number. Nowhere does the rule equate the patentee's awareness and knowledge with all information existing in every possible publicly available source that theoretically could be found.

Case law from various courts makes clear that discovery is not limited by the list of accused devices identified in the infringement contentions and no bright-line

rule limits discovery to only those products specifically accused in a party's infringement contentions. *See, e.g.*, *O2 Micro*, 467 F.3d at 1363; *Honeywell Int'l Inc. v. Acer Am. Corp.*, 655 F.Supp.2d 650, 655-656 (E.D. Tex. 2009). Instead, the rule provides: "Discovery concerning products not explicitly listed in the infringement contentions is appropriate when: 1) the infringement contentions give notice of a specific theory of infringement; and 2) the products for which a plaintiff seeks discovery operate in a manner reasonably similar to that theory." *Nidec*, 2009 WL 3673253 at *2.

2. Keranos's Infringement Contentions Provided Defendants with Adequate Notice of Infringement Theories and Identification of Accused Products.

a) Keranos identified the accused products by the names known to Keranos at the time.

In its IFCs, Keranos identified the accused products by the trade name SuperFlash® and the specific names that the defendants themselves use uniformly and ubiquitously to designate the design architecture employed in the products. "SuperFlash®" refers to a narrow group of products among integrated circuit products with flash memory; the design names refer to very specific SST-originated SuperFlash® memory cell structures. In addition, with the exception of

not listing by individual name all SST's products, Keranos identified specific part numbers and product families to the extent known to it.[8]

SST, having developed SuperFlash technology, readily understood from Keranos's IFCs precisely the products Keranos was accusing. The remaining Manufacturing Defendants, having obtained licenses and the know-how to make SuperFlash® memory, [redacted] for the use of the SuperFlash® memory technology (in accordance with contractual obligations identifying products with this designation), also knew from the IFCs which of their products were the subject of Keranos's patent infringement claims. Similarly, the Customer Defendants selected the SuperFlash memory for their products and, further, had access to this information through the Manufacturer Defendants from whom they obtained semiconductor products with SuperFlash memory cells.

Keranos's identification of accused instrumentalities, in other words, was neither over-inclusive nor under-inclusive. Keranos reasonably informed defendants that the accused products include all SuperFlash® products using ESF1, ESF2 and ESF3 designs and ***only*** those SuperFlash products. Keranos's identification of accused instrumentalities, along with its highly detailed claim

[8] Because SST originated 'SuperFlash' and its huge product base is made up mostly, if not entirely, of 'SuperFlash,' Keranos did not separately include all the individual SST SuperFlash product names or families known to it at the time. Instead, it included a representative product number.

charts, were "specific enough to give [defendants] notice of plaintiff's infringement claims" and to provide "adequate notice and information with which to litigate their cases" in accordance with the overriding function of Rule 3-1. Thus, contrary to the district court's finding that Keranos sought to add "thousands of additional products," the amendments to the IFCs did not expand the scope of accused products in any way. Nor did the amendments shift the infringement theories Keranos disclosed in its original IFCs. *See*, *Bender v. Advanced Micro Devices, Inc.*, No. C-09-1149, 2010 WL 363341, at *1 (N.D. Cal. Feb. 1, 2010) ("The rule is also intended to require the party claiming infringement 'to crystallize its theories of the case early in the litigation and to adhere to those theories once disclosed.'").

Tellingly, not a single defendant moved to strike the IFCs or sought clarification from the court. Even in opposing Keranos's motion to amend, no defendant offered any explanation as to how identification of product numbers would have added in any material way to their understanding of the scope or theory of Keranos's infringement claims as set out in the IFCs.

b) Discovery confirmed the sufficiency of Keranos's identification of accused products in the IFCs.

Discovery conducted after Keranos filed its motion to amend confirmed the sufficiency of Keranos's identification of the accused products by the name "SuperFlash" and the specific design architectures. As set out in Section II.I.2

Confidential Material Redacted

above, the Manufacturing Defendants confirmed during discovery that [redacted] In fact, defendants' technical expert report **never mentioned a single individual product number for any of the defendants' products.**

With the exception of SST whose entire EEPROM product base is SuperFlash, Keranos identified in its IFCs numerous specific product names or models of which it was aware and knew at the time it prepared its IFCs, as Rule 3-1 expressly provides. Keranos sought to amend the IFCs as a practical matter

promptly after obtaining the part numbers through discovery to include greater detail regarding the accused products. This procedure has been approved in other cases. For example, in *3Com Corp. v. D-Link Systems, Inc.*, the district court allowed the plaintiff to amend its original infringement contentions to identify additional products not previously listed by product number. In doing so, the court held the amendment was appropriate where the newly listed products "are merely more detailed listings of the originally accused products with substantially similar functionality." *3Com Corp. v. D-Link Systems, Inc.*, No. C-03-2177, 2007 WL 949599, *7 (N.D. Cal. Mar. 27, 2007). The court further observed that amendment may be appropriate "if a defendant is put on notice by the naming of several products within a line of products possessing similar functionality." *Id.* at *8.

In sum, Keranos's identification of the accused products provided defendants with the requisite notice of the products Keranos was accusing (SuperFlash®), and those it was not (products containing other forms of flash memory). The individual product numbers that Keranos sought to include through amendment of its IFCs were "merely more detailed listings of the originally accused products with substantially similar functionality." *3Com Corp.*, 2007 WL 949599 at *7. Ironically, had Keranos not taken upon itself to seek amendment of the IFCs to list the product numbers learned in discovery, the litigation would have proceeded to completion covering all SuperFlash® products. The district court's

order, based on the conclusion that the Keranos's IFCs did not comply with Rule 3-1, is clearly erroneous and should be reversed.

3. Keranos's Identification of Product Family Names Complies with Rule 3-1.

In its IFCs, Keranos identified defendants' accused products by specific product families, *e.g.*, Freescale's HC08 and 68HC11 products, TI's MSP430 family, Microchip's PIC18, PIC24 and PIC32 products. Even if the court were correct in concluding that identification of the accused products by the name "SuperFlash" and the design architecture did not comply with Rule 3-1(b), the court clearly erred by finding that Keranos's identification by product family number or name is inconsistent with the rules.

First, the district court based its ruling on the wholly unsupported conclusion that "Plaintiff made virtually no effort to specifically identify the accused products as required by the local patent rules." A13. To the contrary, the record is undisputed that Keranos made enormous efforts to identify the accused products. *See*, Section II.E. above. The district court's conclusion, flatly unsupported by the evidence, is clearly erroneous.

Second, even SST referenced its products by "family part numbers," such as SST25xx, in responding to Keranos's discovery requesting the identity of SST's accused SuperFlash products. A6223. Given this obvious admission regarding the sufficiency of identifying the accused products by family number, the court clearly

erred in adopting defendants' contrary arguments in their oppositions to the motion to amend. Third, for the most part, Keranos's references to product family number were as specific as possible at the time and represented the product numbers known to Keranos, as required by Rule 3-1.

C. The District Court Abused Its Discretion in Ruling that Keranos Lacked Good Cause to Amend Its Infringement Contentions.

Local Patent Rule 3-6(b) allows a party to amend its infringement contentions upon a showing of good cause: "When determining whether to grant leave to amend, the Court considers: (1) the explanation for failure to meet the deadline (*i.e.*, the moving party's diligence); (2) the importance of the thing that would be excluded; (3) the potential prejudice in allowing the thing that would be excluded; and (4) the availability of a continuance to cure such prejudice." *Nidec*, 2009 WL 3673253 at *1.

1. Keranos Was Diligent in Identifying SuperFlash Products and the District Court's Ruling Otherwise is Clearly Erroneous.

The record does not support the district court's finding that Keranos was not diligent in its search for and identification of defendants' infringing products based on publicly available information. Thus, the court's ruling is clearly erroneous. As set forth in Section I.E. above, Keranos spent hundreds of hours attempting to individually identify defendants' products that utilize SuperFlash® memory or products. Keranos extensively searched defendants' websites, hired experts and

reviewed hundreds of data sheets. While Keranos was able to find public disclosures that defendants made, sold or imported products incorporating SuperFlash technology, in most cases the publicly available information did not designate specific products as including SuperFlash. Most of the defendants admit as much in their opposition memoranda submitted to the district court. *See*, Section I.H.2., above.

Keranos also reversed engineered some products after making an educated guess that they contained SuperFlash based on Keranos's research. But, in some instances, the reverse engineering demonstrated the products did not contain SuperFlash. A1143. Clearly, the Patent Local Rules do not require Keranos to reverse engineer thousands of potentially infringing products containing flash memory to discern whether the flash memory is SuperFlash. "[T]he question of whether [the disclosing party] conducted 'reverse engineering or its equivalent' is not synonymous with whether it has complied with Patent [Rule] 3-1, which, as discussed, requires only to set forth specific theories of infringement. *STMicroelectronics, Inc. v. Motorola Inc.*, 308 F.Supp.2d 754, 755 (E.D. Tex. 2004) (*citing Network Caching Technology, LLC v. Novell, Inc*, No. C-01-2079, 2003 WL 21699799, *4-5 (N.D. Cal. Mar. 21, 2003)).

In the district court, defendants argued, and the court seems to have agreed, that Keranos should have indiscriminately accused thousands of products as

containing infringing SuperFlash based only on datasheets that indicated the products generically contained "flash" memory. While Keranos was able to accuse a handful of such products based on its investigation, Keranos would have been reckless to accuse all of defendants' flash memory products of infringing, particularly where Keranos's experience with reverse engineering taught Keranos otherwise. In effect, the court's ruling penalizes Keranos's reasonable and careful behavior in its infringement accusations, which is inconsistent with the goals of the discovery rules.

Further, the district court clearly erred when it uniformly applied its finding that Keranos lacked diligence in searching for and naming products of ***all*** defendants. As set forth below, the availability of public information regarding the products was far from uniform among defendants. Yet, the court applied its findings universally even where the record includes no support that Keranos could have found the product information from public sources.

a) The district court's finding that Keranos lacked diligence in identifying Samsung's accused products is clearly erroneous.

The district court's ruling regarding Samsung is unsupported by any evidence. Tellingly, Samsung did not file an opposition to Keranos's motion to amend. The court cited no evidence that Keranos could have discovered specific Samsung product names. Under Eastern District of Texas Local Civil Rule 7(d),

Samsung's failure to oppose the motion required the district court to find in Keranos's favor. *See* E.D. Tex. L. Civ. R. 7(d) ("In the event a party fails to oppose a motion in the manner prescribed herein, the court will assume that the party has no opposition.").

b) The district court's finding that Keranos lacked diligence in identifying TSMC's accused products is clearly erroneous.

TSMC admitted in its opposition to Keranos's motion that Keranos could ***not*** have found TSMC's product numbers from public information and no exercise of diligence by Keranos would have uncovered that information. Even the joint opposition memorandum confirmed this to the district court: "Keranos could have identified virtually all of the products it seeks to add had it exercised diligence in the two years after it filed the complaint (**with TSMC as an exception**)." A1819 (emphasis added). Indeed, TSMC's own witnesses' testimony confirmed that Keranos could not have found TSMC's internal part numbers from public sources. A15516 (81:15-23). Moreover, Keranos properly referred to TSMC's own tradename – "EMBFLASH" – for its version of the SuperFlash memory cells. Thus, the district court's ruling with respect to TSMC was clearly erroneous.

c) No evidence supports the district court's conclusion that Keranos could have identified from public sources the majority of part numbers for other defendants.

In opposition to the motion to amend, Microchip, ADI, and TI conceded that public information merely identified their semiconductor devices as including generic "flash memory," but did not mention "SuperFlash" at all. Freescale made the same concession regarding a majority of its products. "Flash memory" is ***not*** synonymous with "SuperFlash memory," but rather is a generic designation for a generic type of non-volatile memory that can be electrically erased and reprogrammed. A1142-3.

As discussed above, Keranos would have been utterly reckless had it accused all of defendants' products with flash memory of infringing – on the mere chance that the products included SuperFlash.

As with the Manufacturer Defendants, it was a practical impossibility for Keranos to specifically identify every product name/number for every product containing SuperFlash memory from every Customer Defendant. Indeed, most of the Customer Defendants point to no publicly available information that identified their specific products both by individual product numbers and by reference to SuperFlash.

Confidential Material Redacted

Because the district court held Keranos lacked diligence for failing to behave irresponsibly, the court's order is clearly unreasonable, unsupported by any evidence and should be reversed.

d) Publicly available information, even when it referenced "SuperFlash," was often not sufficient to allege infringement.

As Keranos explained to the district court, Keranos's infringement theories and damages calculations only included products from 2004 to 2008. Little in the pre-filing datasheets or Internet searches Keranos obtained in 2010 could categorically confirm that those products were available during that earlier period and used SuperFlash. Keranos provided the district court with several examples to support its position that defendants' "public" datasheets alone were an unreliable source to identify accused products:

- After the close of discovery, Freescale's attorney advised Keranos's attorney that many products Freescale identified in discovery as containing SuperFlash were actually made by other technologies [A6594]; yet, Freescale's datasheets identify those products as containing SuperFlash. *See*, *e.g.*, datasheet for Freescale's MC8HC908JB8 [A6604-5].
- [redacted]

Confidential Material Redacted

These concessions support Keranos's contention that public information was not sufficiently conclusive to identify infringing products during the relevant period.

2. The Remaining Factors Heavily Favor Keranos's Proposed Amendment, Particularly in Light of the New Evidence Produced in Discovery.

The remaining pertinent factors to be considered to determine good cause to amend infringement contentions weigh ***exclusively*** in favor of Keranos, including: (i) the importance of the thing excluded; (ii) the potential prejudice in allowing the thing that would be excluded; and (iii) the availability of a continuance to cure such prejudice. The district court's failure to consider these factors, which utterly eclipse the court's finding of lack of diligence, resulted in a clearly unreasonable and erroneous decision.

a) The products excluded were critically important to Keranos's claims.

As described above, the district court's order prohibited Keranos from pursuing its patent infringement claims against nearly all the accused SuperFlash products, and ultimately resulted in judgment in favor of defendants, even as discovery was completed and trial preparations were underway to adjudicate all of such products. Instead of Keranos maintaining claims for damages in excess of $20 million against the Manufacturer Defendants alone [A15625, n.4],[9] the district

[9] [redacted] A6378.

court's order reduced Keranos claims to a *de minimis* $800 in damages against Microchip and a few tens of thousands of dollars in damages against the Customer Defendants combined. As Keranos explained to the district court, Keranos cannot bring claims against the excluded products in a new case against defendants because the patents-in-suit have expired and all damages are time-barred under 35 U.S.C. § 286. A6378. Accordingly, the importance of the items excluded cannot be overstated.

Instructive here is *Alexsam, Inc. v. IDT Corp.*, No. 2:07-cv-420-CE, 2011 WL 108725 (E.D. Tex. Jan 12, 2011), in which the district court granted the plaintiff's motion to amend its infringement contentions to add new claims and products. In allowing the amendment, the court observed that a denial would have "a significant impact" on the scope of the plaintiff's case because the plaintiff would never be able to seek damages against the excluded products in the future. *Id*. at *2. Here, the district court, rather than avoid that same unnecessarily harsh penalty, entered the order effectively barring Keranos from ever pursuing nearly all of its infringement claims against defendants.

b) Defendants would have suffered no prejudice if the products were included.

The record is undisputed that allowing Keranos's amendment to the IFCs would have caused ***no*** prejudice to defendants. The court's order, made out of the blue and so late in the proceedings, was an unexpected and undeserved windfall to

defendants. At the time Keranos moved to amend, the parties had conducted only a small amount of written discovery, and that discovery covered all the products listed in the amended IFCs. The parties had conducted no depositions, and no experts had been disclosed nor expert opinions rendered; and the *Markman* hearing was scheduled for six months in the future.[10] Defendants had plenty of time to make adjustments to their defense strategies, but that was unnecessary because, [redacted] *See*, Section II.I.2. In addition, the parties spent hundreds of thousands of dollars and thousands of hours preparing and litigating the Manufacturer Case, right up to the point of the pretrial conference, based on claims against all the SuperFlash products Keranos listed in its proposed amended IFCs.

In opposing Keranos's motion to amend, defendants did not state what, if anything, they would have done differently had Keranos's original IFCs listed the specific product numbers. Nor did they provide an explanation as to how their legal theories or trial strategies would have changed based on the amendments. In

[10] During the first eighteen months following Keranos's filing of the complaint, the case did not proceed mostly due to the significant motion practice and venue challenges brought by defendants.

other words, defendants demonstrated no prejudice that the amendment would have allegedly caused.

c) No continuance was needed because there was no prejudice to cure.

As noted by many courts, when prejudice is lacking, consideration of a continuance is unnecessary. *See, e.g.*, *Coopervision, Inc. v. CIBA*, 480 F.Supp.2d 884, 890 (E.D. Tex. 2007). Here, as demonstrated above, defendants would have suffered no prejudice if the district court had timely granted leave to amend. In any case, no continuance would have been necessary because, at the time of Keranos's motion, no date for trial had been set and even the *Markman* hearing was months in the future. Similarly, no continuance would have been necessary at the time the court issued its order since parties completed the work in the interim.

As with the other two factors the court failed to consider, this factor weighs ***entirely*** in favor of Keranos. Although in certain cases it might be appropriate for the district court to curtail its "good faith" analysis after finding a lack of diligence, the district court's decision to focus exclusively on a lack of diligence (which was factually unsupported) resulted in a drastic sanction against Keranos in the absence of ***any*** prejudice to defendants. The district court's decision was clearly unreasonable.

D. The Court Abused Its Discretion by Imposing a Drastic "Death Penalty" Sanction on Keranos that is Both Unjust and Inconsistent with the Federal Rules of Civil Procedure.

Even if the district court were correct that Keranos's original IFCs did not comply with Rule 3-1 because Keranos was not sufficiently diligent, the district court abused its discretion by imposing a belated, sweeping sanction that effectively dismissed with prejudice almost all of Keranos's claims against all defendants. The district court improperly imposed the most drastic sanction available for Keranos's purported non-compliance with Rule 3-1(b), without: (1) issuing notice or warning and providing Keranos an opportunity to cure; (2) finding bad faith or willfulness; (3) finding defendants were prejudiced; or (4) considering a less drastic remedy or sanction. As set forth below, the district court's order is both unjust and inconsistent with the Federal Rules of Civil Procedure and therefore should be reversed.

1. Enforcement of Local Patent Rules Must Not Conflict With the Federal Rules of Civil Procedure.

Courts have variously described the Patent Local Rules both as case management/scheduling orders and discovery devices. *See*, *e.g., O2 Micro Int'l Ltd. v. Monolithic Power Sys., Inc.*, 467 F.3d 1355, 1365 (Fed. Cir. 2006) ("the [patent local] rules are essentially a series of case management orders"); *Finisar Corp. v. DirecTV Group*, 424 F.Supp.2d 896, 899 (E.D. Tex. 2006) (referring to the Patent Local Rules as exercise of court's authority to manage docket, including

"discovery provisions"); *Samsung SDI Co., LTD. v. Matsushita Electric Industrial Co., Ltd.*, No. CV05-8493, 2006 WL 5097360, at *1 (C.D. Cal. June 5, 2006) (observing infringement contentions are "essentially a 'discovery device'" to streamline discovery process) (internal citations omitted).

Regardless of how they are characterized, "[t]o be valid, local rules must be consistent with both acts of Congress and the Federal Rules of Civil Procedure." *O2 Micro*, 467 F.3d at 1365. "Interpretation and enforcement of the discovery provisions of local rules should not conflict with, and should harmonize with, the discovery provisions of the Federal Rules of Civil Procedure." *Finisar*, 424 F.Supp.2d at 899. This Court has recognized that "[i]t is foreseeable that a local patent rule could conflict with the spirit, if not the letter, of the broad discovery regime under the Federal Rules of Civil Procedure, especially given the particular importance of discovery in complex patent cases." *O2 Micro*, 467 F.3d at 1365. "The local patent rules [must] yield to the Federal Rules of Civil Procedure in the event of a direct conflict. *Id*. at n.11.

2. A Drastic "Death Penalty" Sanction Is Appropriate Only In Extreme Circumstances and Only as a Remedy of Last Resort.

The Patent Local Rules do not specify actions the district court may or must take if a party does not comply with Rule 3-1. Any sanction, however, must be just. *Id*. at 1363, citing Fed. R. Civ. P. 16(f) and 37(b)(2)(B). The Federal Rules

of Civil Procedure authorize sanctions both when a party fails to comply with scheduling orders or commits discovery violations (s*ee* Fed. R. Civ. P. 16(f), 37(b) and (c)), but again, any such sanction must be just. *Insurance Corp. of Ireland, Ltd. v. Copagnie des Bauxites de Guinee*, 456 U.S. 694, 707, 102 S.Ct. 2099 (1982); *Compaq Computer Corp. v. Ergonome Inc.,* 387 F.3d 403, 413 (5th Cir. 2004). To be both valid and just, sanctions imposed for non-compliance with a patent local rule must not be inconsistent with those sanctions allowed under the Federal Rules of Civil Procedure.

In its most drastic application, Fed. R. Civ. P. 37 ("Rule 37") permits a court to strike claims from the pleadings and even to dismiss the action or render a judgment by default against the disobedient party. *Roadway Express, Inc. v. Piper,* 447 U.S. 752, 763, 100 S.Ct. 2455 (1980) (*quoting* Fed. R. Civ. P. 37(b)(2)). However, as the Fifth Circuit has repeatedly emphasized, the court's discretion "is not unlimited," and "a dismissal with prejudice is a 'draconian' remedy, or a 'remedy of last resort' only to be applied in extreme circumstances." *Batson v. Neal Spelce Associates, Inc.,* 765 F.2d 511, 515 (5th Cir. 1985); *accord*, *Refac Intern., Ltd. v. Hitachi, Ltd.*, 921 F.2d 1247, 1254 (Fed. Cir. 1990) (citing 9th Cir. law) ("[s]evere sanctions such as taking allegations as established and awarding judgment on that basis, dismissal and default judgment are authorized only in extreme circumstances.") (citing 9th Cir. law); *see also, Bender v. Maxim*

Integrated Products, Inc., No. C-09-01152, 2010 WL 2991257, *2 (N.D. Cal. Jul. 29, 2010) ("*Bender II*") (denying motion to dismiss even where patentee failed to comply with Patent Local Rule 3-1 for third time).

First and foremost, to warrant imposition of these extreme sanctions, the sanction must be based on a finding of bad faith or willful misconduct *Smith & Fuller, P.A. v. Cooper Tire & Rubber Com*, 685 F.3d 486, 488 (5th Cir. 2012); *Batson*, 765 F.2d at 514; *Refac Int'l,* 921 F.2d at 1254. Second, dismissal is not proper if the deterrent value can be substantially achieved by use of less drastic sanctions. *Batson*, 765 F.2d at 514; *Refac Int'l*, 921 F.2d at 1254 ("The district court must take into account … the availability of less drastic sanctions."); *Pressey v. Patterson*, 898 F.2d 1018, 1021 (5th Cir. 1990); *Compaq Computer Corp.,* 387 F.3d at 413. Third, "[t]he non-sanctioned party's trial preparation also must be substantially prejudiced by the violation." *Batson* 765 F.2d at 514; *Refac Int'l*, 921 F.2d at 1254. Finally, dismissal may be inappropriate where the fault is not of the party's or when a party honestly misunderstood the court's instructions. *Batson* 765 F.2d at 514; *Refac Int'l*, 921 F.2d at 1254.

Instructive here is *ClearValue, Inc. v. Pearl River Polymers, Inc.*, 560 F.3d 1291 (Fed. Cir. 2009). In *ClearValue*, the plaintiffs intentionally withheld from discovery critical test results considered by their testifying expert and then provided misleading testimony about their handling of the test results. Discovery

of the withheld test results occurred on the third day of trial, after which the district court held a sanctions hearing. *Id.* at 1294; s*ee*, *ClearValue, Inc. v. Pearl River Polymers, Inc.*, 242 F.R.D. 362, 367 (E.D. Tex. 2007) ("*ClearValue Sanctions Decision*"). At the sanctions hearing, the district court "determined the misconduct was 'egregious' and 'extremely prejudicial' to [the defendant] at the late stage in the case and that drastic sanctions were appropriate." *ClearValue Sanctions Decision*, 242 F.R.D. at 372. The district court struck the plaintiff's pleadings, entered judgment for defendant, and awarded fees and costs as sanctions. *Id*.

On appeal, this Court agreed with the district court that the plaintiffs had violated Rule 26 and the court's discovery order. *ClearValue*, 560 F.3d at 1303. This Court did not agree, however, that dismissal of the case was proper, and held the district court abused its discretion for doing so. *Id*. at 1308. The Court applied the four factors the Fifth Circuit considers in determining whether dismissal is the appropriate discovery sanction. In that case, the Court agreed with the district court that the plaintiffs acted in bad faith, defendants were prejudiced by plaintiffs' actions, and that sanctions were appropriate. *Id*. at 1308. Notwithstanding these critical findings (not present here), the Court reversed the district court's dismissal order, holding that the plaintiffs' conduct, ***although egregious***, did not warrant dismissal. Moreover, the Court noted that while plaintiffs' discovery misconduct was sanctionable, the conduct was less egregious than conduct in other cases where

the Fifth Circuit reversed orders of dismissal for discovery violations where lesser sanctions were available.

Also instructive is *Bender II*, *supra*. In *Bender II*, the plaintiff served original, first, and second amended infringement contentions, none of which complied with Rule 3-1. The first amendment followed agreement by the parties, while the second amendment followed an order of the court. *Bender II*, 2010 WL 2991257, *1 (discussing case history). The defendant moved for the sanction of dismissal on the ground the second amended infringement contentions ("SICs") failed to comply with Rule 3-1 and the court's order requiring plaintiff to amend the second time. *Id*. Despite finding the SICs failed to comply with the court's order, the court refused to dismiss the action because plaintiff's failure to comply was not done willfully or in bad faith. *Id*. at *5. Instead, finding defendant would not be prejudiced by denying dismissal, the court delayed requiring defendant to produce certain discovery until plaintiff complied with Rule 3-1. *Id*.

As discussed below, Keranos's purported non-compliance with Rule 3-1 involved no willfulness or bad faith and did not prejudice defendants in any way. In addition, a lesser sanction, if any, is certainly available. The court's last minute drastic sanction for Keranos's actions is flatly inconsistent with the Federal Rules of Civil Procedure and should be reversed.

3. No Relevant Factors Supported Dismissal.

a) The district court made no finding – and no evidence exists – of willfulness or bad faith.

Keranos and defendants both apprised the district court that its order would "have a devastating impact on Keranos's case," and would deprive Keranos of virtually all of its claims. A6378; A6543. Notwithstanding that knowledge, the district court imposed its severe sanction at the last minute without finding bad faith by Keranos.

In fact, the record demonstrates substantial good faith by Keranos. Keranos conducted hundreds of hours of pre-litigation investigation regarding the accused products. It then timely served its IFCs in accordance with the scheduling order. The IFCs included detailed claim charts mapping the claim elements to structures of the SuperFlash memory cells shown in annotated figures and electron micrographs. Keranos promptly sought discovery regarding product numbers as soon as discovery opened, and it further demonstrated ***good faith*** by, on its own initiative, seeking to amend the IFCs to include product numbers for the accused products after receiving the information from defendants in discovery – even though Keranos faced no motion to strike its IFCs for being incomplete or otherwise insufficient. Importantly, as Keranos informed the district court, Keranos earnestly believed it complied with Rule 3-1 (and gave defendants adequate notice of its infringement theories) by identifying the accused products by

the name "SuperFlash." A6566-7. These circumstances fall well short of the bad faith or willfulness required to impose a death penalty sanction.

b) Any required deterrent value could be achieved through lesser sanctions.

Permissible sanctions under Rule 37 include many with less impact than the draconian remedy of dismissal. Lesser sanctions may include payment of reasonable expenses caused by the party's failure to comply or staying the proceedings until the order is obeyed. *See*, *ClearValue*, 560 F.3d at 1304; Fed. R. Civ. P. 16(f); 37(b)(2)(A)(iv). "[T]he Fifth Circuit 'has consistently held that a district court, when considering the imposition of sanctions for discovery violations … should impose the least severe sanction that will accomplish the desired result.'" *ClearValue*, 560 F.3d at 1305 (*quoting United States v. Garrett*, 238 F.3d 293, 298 (5th Cir. 2000)). Where, as here, Keranos did not act in bad faith and no prejudice befell defendants, the district court should have considered a lesser sanction, such as ordering plaintiff to pay defendants' reasonable costs incurred as a result of providing products that Keranos improperly included in its amended IFCs. But, of course, defendants incurred no such costs.

c) Defendants' trial preparation would not have been prejudiced were Keranos permitted to amend its IFCs.

As discussed above, at the time Keranos sought to amend its IFCs, there would not have been prejudice to defendants had the court timely granted leave.

The lack of prejudice is conclusively demonstrated by the fact that the parties went on to conduct the case, without any interruption, as if all SuperFlash products were included. In sum, defendants' trial preparation, to the extent completed, would have been precisely the same with or without the amendments being allowed. The severe sanctions imposed are unjust in light of the lack of prejudice.

d) At worst, Keranos misunderstood the requirements of Rule 3-1.

As a final factor, dismissal may be inappropriate where the party misunderstood the court's instructions. Here, as Keranos informed the court, it earnestly believed its disclosure was sufficient under Rule 3-1. The record shows that Keranos made significant efforts to comply with Rule 3-1 as Keranos and its attorneys understood it. Keranos plainly did not attempt to withhold its theories from defendants. The circumstances do not present a case where Keranos's level of fault engenders such extreme sanctions.

4. The District Court's Draconian Order Stands in Stark Contrast to Other Orders Entered for Similar Purported Violations.

The district court's draconian order stands in glaring contrast to other orders entered by the Eastern District of Texas and other courts in similar circumstances. For example, in *Alexsam, Inc. v. IDT Corp., supra*, the district court granted the plaintiff's motion to amend its infringement contentions to add new claims and products, notwithstanding finding a lack of diligence. In allowing the amendment,

the district court took into consideration the fact that plaintiff otherwise would never be able to assert the claims or seek damages against the products in the future, and that excluding the amendments would have a significant impact on the scope of plaintiff's claims. In *Bender II*, *supra*, based on the plaintiff's lack of bad faith, the district court denied the defendant's motion for the sanction of dismissal and allowed the plaintiff to amend its contentions again even after the plaintiff ***thrice*** failed to comply with Rule 3-1. In *Oracle Am., Inc. v. Google Inc*., No. C-10-03561, 2011 WL 4479305 (N.D. Cal. Sept. 26, 2011), a case defendants cited below, the district court allowed the patentee to ***twice*** supplement its infringement contentions after the defendant challenged their sufficiency. In fact, the plaintiff had a third opportunity to supplement, but chose not to do so. The court, in *Samsung SDI Co.*, *supra*, denied the defendants' motion to dismiss claims and strike related infringement contentions after the plaintiff amended its infringement contentions three times.

Here, Keranos sought leave to amend once, at a time very early in the litigation. It received the court's notice that its claims were deficient only at the last minute prior to trial and had no opportunity to cure. Consistent with the Federal Rules of Civil Procedure and orders entered in similar cases, Keranos should be permitted to proceed with its claims.

5. The District Court's Order Is Particularly Harsh for Keranos Given the District Court's Delay in Ruling and the Significant Resources Expended by Keranos During the Delay.

Keranos filed its motion to amend when the case was in its very early stage and litigation costs were *relatively* small. The district court, however, did not rule until more than a year had passed from the filing. This long delay resulted in severe penalties to Keranos given the district court's eventual ruling. During that time, Keranos completed discovery in the Manufacturer Case as to all SuperFlash® products Keranos identified in its proposed amended IFCs. The parties conducted thirty depositions, including in Taiwan and Korea and various locations in the United States. Keranos prepared and responded to numerous discovery requests and reviewed voluminous documents produced by defendants. Keranos prepared claim construction briefing on the three patents-in-suit and appeared for the *Markman* hearing. The parties exchanged opening and rebuttal technical and damages expert reports and took expert depositions addressing all the SuperFlash® products. After filing its motion to amend and before the district court ruled, Keranos incurred hundreds of thousands of dollars in costs alone prosecuting its claims against defendants.

While Keranos is mindful of the strains placed on judicial resources, the district court's ultimate drastic ruling coupled with the delay has the enhanced

impact on Keranos in that its claims are dismissed, but only after it expended significant financial resources in prosecuting those claims.

V. CONCLUSION AND RELIEF SOUGHT

For the foregoing reasons, the Court should reverse the judgments in favor of defendants and the orders denying Keranos's motion to amend its Rule 3-1 infringement contentions and remand the cases to the district court.

Respectfully submitted,

AGILITY IP LAW, LLP
149 Commonwealth Dr., Ste. 1033
Menlo Park, CA 94025
650-227-4800

Dated: August 11, 2014 By: */s/ Michelle G. Breit*
James C. Otteson
Michelle G. Breit

Attorneys for
Plaintiff – Appellant Keranos, LLC

ADDENDUM

ADDENDUM

Addendum Table of Contents

Date Filed	Dkt. No.	Case No.[11]	Docket Text	Apx. No.
Required Materials				
2/10/2014	251	017	Final Judgment	A1
8/5/2013	151	017	Order	A2
1/3/2014	226	017	Order Denying Motions to Reconsider [154 and 159]	A11
2/4/2014	248	017	Order Granting Motion for Summary Judgment and Dismissal of Claims re Samsung	A14
2/4/2014	249	017	Order Granting Motion for Summary Judgment and Dismissal of Claims re SST, et al.	A15
5/1/2014	084	018	Final Judgment	A17
1/28/2014	056	018	Order Denying Motion to Amend	A18
4/4/2014	078	018	Memorandum Order Regarding Customer Products	A19
5/1/14	083	018	Order Granting Summary Judgment	A25
			US Patent No. 4,795,719	A247
			US Patent No. 4,868,629	A259
			US Patent No. 5,042,009	A287
Statutory Materials				
			E.D. Tex. Local Civil Rule 7(d)	Add. 1
			E.D. Tex. Rules of Practice for Patent Cases 3-1	Add. 2
			E.D. Tex. Rule of Practice for Patent Cases 3-6	Add. 3
			Fed. R. Civ. P. Rule 16(f)	Add. 4
			Fed. R. Civ. P. Rule 37(b)	Add. 5

[11] "Case No." refers to the District Court case the documents were originally filed under: Case No. 2:13-CV-00017-MHS-RSP (017) or Case 2:13-cv-00018-MHS-RSP (018).

United States District Court
EASTERN DISTRICT OF TEXAS
TYLER DIVISION

KERANOS, LLC	§	
	§	
v.	§	Case No. 2:13-cv-17-MHS-RSP
	§	
SILICON STORAGE TECHNOLOGY, INC., et al.	§	

FINAL JUDGMENT

In accordance with the Court's orders granting summary judgment (Doc. Nos. 248, 249), it is hereby **ORDERED**, **ADJUDGED**, and **DECREED** that Plaintiff's claims be dismissed with prejudice, Defendants' counterclaims be dismissed without prejudice, and final judgment be entered in this case.

All relief not previously granted is hereby **DENIED**.

It is SO ORDERED.

SIGNED this 7th day of February, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

A1

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC	§	
	§	
v.	§	Case No. 2:13-cv-17
	§	
SILICON STORAGE TECHNOLOGY, INC., et al.	§	

ORDER ON PENDING MOTIONS

Plaintiff in this patent infringement action alleges that Defendants—manufacturers of flash memory products—infringe three related patents. (A related case, *Keranos v. Analog Devices, Inc.*, 2:13-cv-18, involves similar claims against customer defendants.) Defendants challenge the patents as invalid.

This order addresses a number of issues pending before the Court.

Motion to Strike Plaintiff's Second Set of Interrogatories (Doc. No. 7)

Defendants ask the Court to strike Plaintiff's second set of interrogatories as exceeding the number allowed by the Court. Considering the briefing and the applicable law, the motion is DENIED.

Defendants argue that Plaintiff agreed to limit the number of interrogatories to 25 per Defendant. Defendants also argue that Plaintiff's first set of interrogatories exhausted this limit based on the rules for counting distinct subparts as separate interrogatories. Accordingly, Defendants ask the Court to strike Plaintiff's second set of interrogatories. Plaintiff counters that Defendant misapplies the standard for counting interrogatories. Plaintiff emphasizes that under

Defendants' application of the rules, their own interrogatories would exceed the limit.[1]

Plaintiff's first set of interrogatories served on Defendant SST contained nine separately numbered interrogatories. Plaintiff's first set of interrogatories as to all remaining Defendants contained seven separately numbered interrogatories. According to Defendants, each subpart should be counted separately based on (1) the number of patents the interrogatory covers; (2) the number of accused products addressed by the interrogatory; and (3) because some interrogatories seek information about facts, legal bases, identification of documents, and the identification of people with relevant information. This case involves three patents and dozens of products manufactured by Defendants. Under Defendants' calculation, Plaintiff's first set of interrogatories amounted to several hundred interrogatories in total, with at least 33 served on each Defendant.

According to the Advisory Committee Notes to Rule 33, parties cannot circumvent the limit on the number of interrogatories "through the device of joining as 'subparts' questions that seek information about discrete separate subjects." The Advisory Committee does not define "subpart," but notes that "a question asking about communications of a particular type should be treated as a single interrogatory even though it requests that the time, place, persons present, and contents be stated separately for each such communication." Courts have generally determined that where the parent question can be answered without reference to the subpart, then the subpart should be counted as a distinct interrogatory. *FTC v. Think All Pub, L.L.C.*, No. 4:07-cv-011,

[1] Plaintiff also argues that Defendants waived this objection when they failed to raise it in response to Plaintiff's first set of interrogatories. Also, based on the complicated procedural history of this case, which originated as five separate actions that were realigned by the Court in December 2012, Plaintiff contends the parties agreed to 55 interrogatories per Defendant. Because the Court finds Defendants' argument without merit, the Court does not reach Plaintiff's alternative arguments.

2008 WL 687455, at *1 (E.D. Tex. Mar. 11, 2008).

Under this standard, the number of patents or accused products does not present a blanket basis for double-counting Plaintiff's interrogatories. The three patents in this action all cover similar technology, specifically split-gate flash memory, and the accused products incorporate similar or in some cases identical technology. Accordingly, Plaintiff's interrogatories should not be counted multiple times based on the number of patents or accused products responsive to each interrogatory.

Furthermore, as a general rule, Plaintiff's requests seeking facts, documents, *and* identification of persons with information do not automatically amount to three separate interrogatories. *See, e.g.*, *Stamps.Com, Inc. v. Endicia, Inc.*, No. CV 06-7499-ODW, 2009 WL 2576371, at *3 (C.D. Cal. May 21, 2009) ("[R]equests for facts, persons with knowledge of those facts and documents containing those facts should be considered one interrogatory because they are subsumed within the primary question of facts supporting defendants' infringement and validity contentions."); *but see FTC v. Think All Pub, L.L.C.*, No. 4:07-cv-011, 2008 WL 687455, at *2 (E.D. Tex. Mar. 11, 2008). Defendants have failed to demonstrate that these interrogatories should be separately counted.

Accordingly, Defendant's Motion to Strike Plaintiff Keranos, LLC's Second Set of Interrogatories (Doc. No. 7) is DENIED. To the extent Defendants have not already responded to the interrogatories or the parties have not otherwise resolved this issue, Defendants are ordered to submit their responses within 14 days of this order. The parties are encouraged to reach agreement if a modified response timeline is warranted (i.e. a rolling response).

Plaintiff's Motion to Compel Non-infringement Contentions (Doc. No. 23)

Plaintiff also challenges Defendants' responses to its non-infringement contentions as insufficient and asks the Court to compel Defendants to provide complete responses. Having considered the issues and the procedural posture of this case, the motion is GRANTED.

Defendants object to Plaintiff's request for complete responses, suggesting that Plaintiff is attempting to prematurely compel expert discovery. But the deadline to exchange expert reports has now passed. Thus, the time is ripe for Defendants to fully supplement their non-infringement contentions. Accordingly, to the extent they have not already done so, Defendants are ordered to serve supplemental responses to Plaintiff's non-infringement contention interrogatories within 14 days of this order.

Plaintiff's Motions for Leave to Amend Infringement Contentions (Doc. Nos. 49, 51)

Before the Court are Plaintiff's Motions for Leave to Amend Infringement Contentions (Doc. Nos. 49, 51)[2] Plaintiff seeks to amend its infringement contentions to add additional products disclosed by Defendants during discovery. For the reasons discussed below, the motions are DENIED.

Local Patent Rule 3-1 requires a party claiming infringement to identify each accused product in its infringement contentions. The "identification shall be as specific as possible," including name and model number, if known. PR 3-1(b). Generally, infringement contentions may only be amended or supplemented upon a showing of good cause. PR 3-6(b). The Court considers four factors when reviewing a motion to amend infringement contentions: "(1) the

[2] Although these motions were recently filed in this case, the Court notes that the motions originally were filed in July 2012. Following the claim construction hearing, the Court denied the motions without prejudice, realigned the parties, and ordered the parties to attend mediation. The motions were refilled pursuant to the Court's order after mediation proved unsuccessful.

explanation for the party's failure to meet the deadline, (2) the importance of what the Court is excluding, (3) the potential prejudice if the Court allows the thing that would be excluded, and (4) the availability of a continuance to cure such prejudice." *Alexsam Inc. v. IDT Corp.*, No. 2:07-cv-420-CE, 2011 WL 108725, at *1 (E.D. Tex. Jan. 12, 2011). As part of the good cause showing, the party seeking to amend must demonstrate that it was diligent in discovering the additional products and in seeking to amend. *Id.*; *see also West v. Jewelry Innovations, Inc.*, No. C 07-1812, 2008 WL 4532558, at *2 (Oct. 8, 2008) (finding that a party must be diligent in discovering the basis for amendment).

Plaintiff's proposed amendment would add thousands of additional products not specifically disclosed in Plaintiff's original infringement contentions. Each of these products incorporates Defendant SST's SuperFlash technology. Accordingly, Plaintiff argues that each of these products fall within the scope of its original infringement contentions, which identified some specific products and generally "identified" other products sold by Defendants in the United States that incorporated the SuperFlash technology.

Plaintiff also notes that it is not adding any new claims or altering its infringement theory. The products Plaintiff seeks to add were disclosed by Defendants in response to Plaintiff's interrogatories.

Defendants counter that Plaintiff failed to exercise due diligence prior to filing this action or in the 18 months between when Plaintiff first filed this action and when Plaintiff served its original infringement contentions. Specifically, Defendants note that the many of the products were designated as including the SuperFlash technology in documents already in Plaintiff's possession at the time of its original infringement contentions. Furthermore, Defendants note that

other products could have been identified using public documents available on Defendants' websites or using third-party websites to search product datasheets. Additionally, Defendants insist that once they identified the additional products, Plaintiff continued to delay in amending its infringement contentions. Defendants argue that Plaintiff's failure to exercise diligence in identifying infringing products cannot be excused under the facts of this case.

Plaintiff counters that the documents in their possession when they served their original infringement contentions or that were otherwise publically available do not sufficiently disclose that the products used the SST SuperFlash technology. Instead, many of the documents merely noted that the products incorporated flash technology or referred to the products as "microprocessors." Through expensive reverse engineering, Plaintiff claims it found that some of these products did not include the accused SuperFlash technology. Furthermore, Plaintiff notes that none of these documents indicated whether the products were sold or marketed in the United States. Plaintiff concludes that the only feasible way to identify all of the infringing products was through discovery. Plaintiff claims that it exercised due diligence when it promptly served on Defendants interrogatories aimed at identifying the products. But according to Plaintiff, Defendants dragged their feet in responding to the interrogatories. Furthermore, Plaintiff alleges that Defendants cannot claim to be prejudiced by the amendment as each of the new products falls within the general scope of its original infringement contentions.

Plaintiff has failed to demonstrate that it acted diligently in searching for and naming the additional products that incorporate the accused technology. Defendants demonstrate a number of means through which Plaintiff could have identified products incorporating the SuperFlash technology. Instead of making these efforts, Plaintiff chose to list a handful of exemplar products

and then demand that Defendants disclose additional products. But the burden is on Plaintiff, not Defendants, to search for and identify infringing products to the extent possible based on publically available information. *Am. Video Graphics, L.P. v. Elec. Arts, Inc.*, 359 F. Supp. 2d 558, 560 (E.D. Tex. 2005) ("The Patent Rules demonstrate high expectations as to plaintiffs' preparedness before bringing suit, requiring plaintiffs to disclose their preliminary infringement contentions before discovery has even begun.").

The Court finds unavailing Plaintiff's argument that the publically available documents did not conclusively demonstrate that the products incorporated the accused technology or address whether the products were sold in the United States. First, many of the products identified by Plaintiff in its original contentions were identified using the exact or similar documents that Plaintiff now claims are insufficient. Plaintiff simply did not exhaust its search, instead choosing to shift the burden to Defendants.

Furthermore, the Court finds significant the sheer volume of the additional products Plaintiff seeks to add. It may be conceivable that even acting diligently, Plaintiff would have failed to identify some of the products that incorporate the accused technology. Instead, Plaintiff found the efforts to identify products to be expensive and cumbersome and instead disclosed only a few products in its original contentions, then demanded that Defendants identify the remaining products that incorporate the accused technology. This is contrary to Plaintiff's responsibility under the local patent rules and demonstrates a lack of diligence by Plaintiff.

Because Plaintiff has failed to demonstrate it acted diligently, Plaintiff's Motions for Leave to Amend Infringement Contentions (Doc. Nos. 49, 51) are DENIED.

Plaintiff's Motion to Limit Defendants' Asserted Prior Art References (Doc. No. 18)

Also before the Court is Plaintiff's motion to limit defendant's asserted prior art references (Doc. No. 18). Having considered the parties' arguments and the applicable law, the motion is GRANTED.

Defendants have disclosed 93 prior art references related to the 10 asserted claims. Plaintiff asks the court to limit Defendants to two to three references per claim. Plaintiff also asks that Defendants be ordered to amend their invalidity contentions to reflect this limitation.

Defendants argue that Plaintiff's request is premature and seeks to prevent Defendants from asserting combinations of more than three references. The Court disagrees.

To the extent Defendants argue that Plaintiff's request is premature, the Court notes that discovery has substantially completed, thus alleviating Defendants' concerns.[3] Additionally, Defendants' concern about limiting combinations is misplaced. The Court routinely limits Defendants to two or three prior art references. *See, e.g., Global Session LP v. Travelocity.com LP*, No. 6:10-cv-671-LED-JDL, slip op. at 1 (E.D. Tex. November 14, 2012), ECF No. 415. The Court will count a combination as a single prior art reference.

[3] The Court also notes that the Advisory Council for the United States Court of Appeals for the Federal Circuit recently released a model order for use by district courts. Advisory Council for the United States Court of Appeals for the Federal Circuit, A Model Order Limiting Excess Patent Claims and Prior Art (2013), *available at https://www.docketnavigator.com/images/FinalModelOrderLimitingExcessPatentClaimsAndPriorArt.pdf.* The model order requires the parties to pare down the number of asserted claims and prior art references in two stages: prior to claim construction and again after claim construction. *Id.* In both instances, the parties must limit their

Accordingly, Defendants are order to limit their prior art references to two to three per claim. Defendants must serve Plaintiff with amended invalidity contentions reflecting these limitations within 14 days of this order.

It is SO ORDERED.

SIGNED this 2nd day of August, 2013.

_____
MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

asserted claims or prior art references before the close of discovery. *Id.*

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC	§	
	§	
v.	§	Case No. 2:13-cv-17
	§	
SILICON STORAGE TECHNOLOGY, INC., et al.	§	

ORDER ON MOTIONS TO RECONSIDER

Plaintiff in this patent infringement action alleges that Defendants—manufacturers of flash memory products—infringe three related patents. (A related case, *Keranos v. Analog Devices, Inc.*, 2:13-cv-18, involves similar claims against customer defendants.) Defendants challenge the patents as invalid.

The Court previously ruled on several pending motions, including Plaintiff's request to amend its infringement contentions and Plaintiff's motion to compel Defendants' non-infringement contentions. The Court denied Plaintiff's motion for leave to amend its infringement contentions and granted Plaintiff's motions to compel answers to its interrogatories. Plaintiff now seeks reconsideration of both issues (Doc. Nos. 154, 159). For the reasons discussed below, the motions to reconsider are DENIED.

Reconsideration of Court's Order Compelling Non-Infringement Contentions

In its original motion, Plaintiff challenged Defendants' responses to its non-infringement contentions as insufficient and asked the Court to compel complete responses from Defendants. Defendants objected to Plaintiff's request, arguing that their response was sufficient and that Plaintiff was attempting to prematurely compel expert discovery. The Court noted that the

deadline to exchange expert reports had passed and granted the motion to compel Defendants to respond "to the extent they [had] not already done so."

Plaintiff now argues that the order—which issued after the close of fact discovery—allows Defendants to proceed on facts not timely disclosed. But the Court's previous order compelling Defendants' supplemental non-infringement contentions did not address whether Defendant may present those theories at trial. That issue is the subject of a letter brief filed by Plaintiff (Doc. No. 94) and will be addressed accordingly. Therefore, the Court finds no reason to amend its previous ruling that Defendant respond to Plaintiff's interrogatories, and the motion to reconsider (Doc. No. 154) is DENIED.[1]

Reconsideration of Court's Order Denying Leave to Amend Infringement Contentions

Also, the Court previously denied Plaintiff's request to amend its infringement contentions to add thousands of additional products. The Court found that Plaintiff did not act diligently when disclosing its original infringement contentions and thus did not demonstrate good cause for the amendment.

Plaintiff's motion to reconsider raises the same arguments previously addressed by the Court and does not justify amendment of the Court's previous ruling. Accordingly, the motion to reconsider (Doc. No. 159) is DENIED.

But Plaintiff also seeks clarification of the Court's earlier order. In its original infringement contentions, Plaintiff identified several product families, such as Freescale's HC08 product family. The parties now question whether these products were sufficiently disclosed in

[1] The Court denies the motion without the benefit of Defendants' response, which was not filed. The Court considered Plaintiff's motion and reply and Defendants' surreply.

Plaintiff's original contentions or whether the Court's refusal to allow Plaintiff to amend excludes those products from this action. As discussed in the Court's previous order, Plaintiff made virtually no effort to specifically identify the accused products as required by the local patent rules. Instead, Plaintiff demanded that Defendants identify the allegedly offending products. While the Court does not necessarily foreclose any categorical infringement contentions, Plaintiff's disclosure of entire product families without limited or no attempt to narrow the accused product descriptions is inconsistent with the rules of this court.

It is SO ORDERED.

SIGNED this 3rd day of January, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC, ET AL., PLAINTIFFS, VS. SILICON STORAGE TECHNOLOGY, INC., ET AL., DEFENDANTS, COUNTER-CLAIM PLAINTIFFS, VS. KERANOS, LLC, ET AL., COUNTER-CLAIM DEFENDANTS.	§ § § § § § § § § § § § § § § §	No. 2:13-CV-00017 (MHS)

ORDER GRANTING SUMMARY JUDGMENT AND DISMISSAL OF CLAIMS

IT IS HEREBY ORDERED, and Plaintiff Keranos, LLC's ("Keranos") agrees, subject to its rights of appeal, that summary judgment dismissing Keranos's claims against Samsung Semiconductor, Inc., and Samsung Electronics Co., Ltd. ("Samsung"), with prejudice, is proper and that all claims that have been brought and/or that could have been brought by Keranos against Samsung are dismissed with prejudice.

IT IS FURTHER ORDERED that Samsung's declaratory judgment claims and counterclaims of non-infringement and invalidity of the Patents-in-Suit are dismissed without prejudice.

IT IS SO ORDERED.

SIGNED this 4th day of February, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

IN THE UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC, Plaintiff, *vs.* SILICON STORAGE TECHNOLOGY, INC., et al., Defendants.	Case No. 2:13-cv-00017-MHS-RSP

ORDER

THE COURT, having denied plaintiff Keranos' Motion for Leave to Amend Infringement Contentions [Order on Pending Motions, Dkt. 151] and Keranos' Motion to Reconsider [Order on Motions to Reconsider, Dkt. 226], and having concluded at the Hearing on January 29, 2014 that Samsung's products are excluded, thereby concluding that Keranos is prohibited from proceeding on any patent infringement claims under U.S. Patent Nos. 4,868,629, 4,795,719, and 5,042,009 ("the Patents-in-Suit") against Defendants Silicon Storage Technology, Inc. ("SST"), Microchip Technology, Inc., Freescale Semiconductor Inc., Taiwan Semiconductor Manufacturing Co., Ltd. and TSMC North America (all collectively "Defendants") based on any products other than Microchip's PIC32MX440F128L and SST's 89F58 (hereinafter, the "Two Remaining Products");

FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against those Two Remaining Products;

AND FURTHER, Defendants having agreed to dismiss without prejudice their declaratory judgment claims and counterclaims of non-infringement and invalidity of the Patents-in-Suit; and SST having agreed to dismiss without prejudice its state law tort claims for tortious interference with existing contracts, tortious interference with prospective business relationships and civil conspiracy (collectively "SST State Law Tort Claims") subject to a tolling order;

IT IS HEREBY ORDERED that Keranos' patent infringement claims against the Two Remaining Products are dismissed with prejudice.

IT IS FURTHER ORDERED that summary judgment of all remaining claims of infringement of the Patents-in-Suit against Defendants is granted, and these claims are dismissed with prejudice subject to Keranos' right to appeal.

IT IS FURTHER ORDERED that Defendants' declaratory judgment claims and counterclaims of non-infringement and invalidity of the Patents-in-Suit are dismissed without prejudice.

IT IS FURTHER ORDERED that SST's State Law Tort Claims are hereby dismissed without prejudice.

IT IS FURTHER ORDERED, and the parties have agreed, that SST may reinstate the SST State Law Tort Claims within ninety (90) days following the final disposition of any appeal of, or the expiration of any time period providing an opportunity to appeal, a judgment or order of the Court entered in this action or appellate decision arising out of a judgment or order of the Court entered in this action, whichever is later.

IT IS FURTHER ORDERED, and the parties agree, that any applicable statute of limitations or other applicable law or rule that relates to the time period for filing the SST State Law Tort Claims is tolled until the deadline set out above for reinstatement of the SST State Law Tort Claims. The tolling of the limitations period as set forth in this Order is valid and effective under applicable law and not contingent upon the outcome of any appeal arising out of this action.

IT IS SO ORDERED.

SIGNED this 4th day of February, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC	§	
	§	
v.	§	Case No. 2:13-cv-18
	§	
ANALOG DEVICES, INC.	§	

FINAL JUDGMENT

In accordance with the Court's order dismissing the claims in this case, it is hereby **ORDERED, ADJUDGED, and DECREED** that final judgment be entered in this case.

All relief not previously granted is hereby DENIED.

It is SO ORDERED.

SIGNED this 1st day of May, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC,	§	
	§	
v.	§	Case No. 2:13-CV-18-MHS-RSP
	§	
ANALOG DEVICES, INC., et al.	§	

ORDER

Before the Court is Defendant Keranos, LLC ("Keranos")'s Motion for Leave to Amend its Infringement Contentions to Add Products Recently Disclosed in Defendants' Discovery Responses (Dkt. No. 24, filed May 5, 2013) and Keranos's Motion for Leave to Amend its Infringement Contentions Recently Disclosed in [Defendants'] Discovery Responses (Dkt. No. 27, filed May 5, 2013.) Both motions were originally filed in Case No. 2:11-cv-0331 and were refiled pursuant to the Court's Order (Dkt. No. 45 in Case No. 2:13-cv-17.) The Court, in Case No. 2:13-cv-17, denied identical motions by Keranos (Dkt. No. 151 in Case No. 2:13-cv-17), Keranos then moved for reconsideration of the Court's order (Dkt. No. 154 in Case No. 2:13-cv-17), and the Court denied Keranos's motion for reconsideration (Dkt. No. 226 in Case No. 2:13-cv-17).

For the reasons assigned by the Court, both in denying the identical motions in Case No. 2:13-cv-17 and in denying Keranos's request for reconsideration in Case No. 2:13-cv-17, Keranos's motions are hereby **DENIED**.

SIGNED this 28th day of January, 2014.

ROY S. PAYNE
UNITED STATES MAGISTRATE JUDGE

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC,	§	
	§	
v.	§	Case No. 2:13-CV-18-MHS-RSP
	§	
ANALOG DEVICES, INC., et al.	§	

MEMORANDUM ORDER

Following a hearing held by the Court, Defendants Apple Inc., Analog Devices, Inc., Intel Corporation, International Business Machines Corporation, National Semiconductor Corporation, NXP Semiconductors USA, Inc. and Texas Instruments, Inc. (collectively, "Defendants") filed a Motion for Procedural Order (Dkt. No. 60, filed January 30, 2014.) Following Defendants' Motion, the Parties jointly moved for an entry of a procedural order (Dkt. No. 61, filed January 31, 2014.) The Court granted the Parties agreed order (Dkt. No. 62, filed February 4, 2014.)

Pursuant to the Court's Order Keranos filed a Memorandum Regarding Disputed Products Remaining in the Case (Dkt. No. 68, filed February 21, 2014) (hereinafter "Keranos Memo."), setting forth the results of the Parties' meet and confer process, as to what accused products remain in the case following the Court's Orders dated August 2, 2013 (Case No. 2:13-cv-00017-MHS-RSP, Dkt. No. 151), January 3, 2014 (Case No. 2:13-cv-00017-MHS-RSP, Dkt. No. 226), and January 28, 2014 (Case No. 2:13-cv-00018-MHS-RSP, Dkt. No. 56) (hereinafter "Court's Prior Orders"). Keranos's Memorandum stated that the following products were agreed or disputed and provided argument as to the products Keranos disputes are in the case:

DEFENDANT	AGREED	DISPUTED
ADI	ADUC814, ADUC824, ADUC831	
Apple	8 GB 2nd Generation	iPhone

	iPod Nano	
IBM		IBM System x3100
NXP	LPC2888, LPC2000, LPC2917/2919/01 PNX0161	
Intel	945 Express chipset 82573E LAN	
TI	TMS470RX1, TIMSC1211Y4	CC2430, CC2510Fx, CC2511Fx

Defendants with disputed products subsequently filed documents containing their argument as to why the products they dispute are out of the case: IBM (Dkt. No. 69, filed February, 21, 2014) (hereinafter "IBM Memo"); Apple (Dkt. No. 70, filed February 21, 2014) (hereinafter "Apple Memo"); and Texas Instruments (Dkt. No. 71, filed February 21, 2014) (hereinafter "TI Memo").

DISCUSSION

1) ADI

The Court accepts the Parties' agreement.

2) Apple

Keranos argues that its identification of Apple's "iPhone" is properly definite under the Court's Prior Orders. Keranos argues that its identification of the iPhone was not a disclosure of an "entire product famil[y]," since Apple only had two models of iPhone in the relevant damages period, though Keranos admits Apple sold multiple versions of each model. (Keranos Memorandum at 2.)

Apple argues that Keranos's identification of the "the 'iPhone' is exactly the type of product family that this Court ruled is not properly accused." (Apple Memo at 1 (emphasis removed).) Apple argues that the identification of the "iPhone" is not "as specific as possible" and that Keranos did not attempt to narrow the accused product description. (*Id.* at 2 (emphasis

removed).) Apple argues that, in the "infringement contentions that the Court denied Keranos leave to serve, Keranos identified five different iPhone models . . . ." and that Keranos stated that Apple incorporated the accused memory into "various models of the iPhone." (*Id.* at 2-3 (emphasis removed).) Apple additionally argues that the no iPhone was ever charted according to P. R. 3-1(c). (*Id.* at 3.) Apple argues that two months before Kearnos filed its infringement contentions Apple had timely produced relevant technical information that would have allowed Keranos to timely identify the specific models Keranos now disputes. (*Id.* at 4.)

For the reasons assigned by the Court in the Court's Prior Orders, the Court hereby finds that the identifier "iPhone" represents an entire product family and is not properly definite under the Court's Prior Orders.

The Court accepts the Parties' agreement as to the 8 GB 2nd Generation iPod Nano.

3) IBM

Keranos argues that its identification of the IBM's "IBM System x3100" is properly definite under the Court's orders. Keranos argues that, at some point in discovery, "IBM asserted 'the only accused product identified by Keranos to date is the IBM System x3100 product'" and that IBM did not assert that the disclosure of the IBM System 3100 failed to identify a specific product." (Keranos Memorandum at 2-3.) Keranos notes that in IBM's objections to Keranos's amended infringement contentions, IBM stated Keranos's "original infringement contentions (and complaint) identified an actual IBM product (identified as 'the IBM System x3100 product.'" (*Id.* at 3.) Keranos argues that "the presence of just two models does not indicate that Keranos' identification of the 'IBM System x3100' constitutes the type of disclosure of an 'entire product family' of which the Court was concerned in its order."

IBM argues that Keranos did not identity products "with specificity, e.g., by individualized product number" and that "[i]t is undisputed that the identified IBM Product

Families are general designators that refer to families or groups of multiple products, not specific individualized products." (IBM Memo at 1.) IBM argues that "Keranos told this Court that it 'interprets the Court's Orders to convey that Keranos' infringement contentions stand only to cover individual SuperFlash products which Keranos identified by their specific individualized product number and no other products are or can be accused in this lawsuit." (*Id.* at 3.) IBM argues that IBM has identified at least 23 "specific products in the IBM System x3100 family." (*Id.*) IBM argues the IBM System x3100 is an equivalent identifier to the Samsung "S3F4xxx" line of products that the Court already found to be excluded. (*Id.*) IBM argues that Keranos had documents in its possession that would have allowed it to identify specific products within the IBM System x3100 family. (*Id.*) IBM argues that Keranos's argument that the relevant damages period aids Keranos in properly identifying products is incorrect under the Court's rules, improperly shifts the burden of identifying accused products, and was previously rejected three times by the Court. (*Id.* at 4-5.)

For the reasons assigned by the Court in the Court's Prior Orders, the Court hereby finds that the identifier "IBM System x3100" represents an entire product family and is not properly definite under the Court's Prior Orders.

4) NXP

The Court accepts the Parties' agreement.

5) Intel

The Court accepts the Parties' agreement.

6) Texas Instruments

Keranos argues that the designations "CC2430," "CC2510Fx," and "CC2511Fx" are designations that TI uses internally. Keranos admits that the CC2430 is actually sold in three different flash versions: CC2430F32, CC2430F64, and CC2430F128, depending on how much

flash memory is onboard (e.g. 32/63/128 KB). Keranos admits the CC2510Fx and CC2511Fx were also sold in three different flash versions: CC2510F8/CC2511F8, CC2510F16/CC2511F16 and CC2510F32/CC2511F32, depending on how much flash memory is onboard. Keranos argues that "memory size is not an element of any of the patents in suit" and that "TI suffered no confusion or lack of notice regarding these accused products . . . ." (Keranos Memo at 5.)

TI argues[1] that Keranos did not identity products "with specificity, e.g., by individualized product number" and that "[i]t is undisputed that the identified TI Product Families are general designators that refer to families or groups of multiple products, not specific individualized products." (TI Memo at 1.) TI argues that "Keranos told this Court that it 'interprets the Court's Orders to convey that Keranos' infringement contentions stand only to cover individual SuperFlash products which Keranos identified by their specific individualized product number and no other products are or can be accused in this lawsuit." (*Id.* at 2.) TI argues that the disputed products—CC2430, CC2510Fx, CC2511Fx[2]—undisputedly refer to families and not to specific individualized product numbers, and that Keranos referred to these numbers as "families" in its contentions. (*Id.* at 2-3.) TI argues the disputed products are equivalent identifiers to the Samsung "S3F4xxx" line of products that the Court already found to be excluded. (*Id.* at 3.) TI argues that Keranos's argument that the relevant damages period aids Keranos in properly identifying products is incorrect under the Court's rules, improperly shifts the burden of identifying accused products, and was previously rejected three times by the Court. (*Id.* at 4-5.)

---

[1] TI's motion states that it incorporates approximately one page of the IBM Memo by reference. Out of respect for the Court's order adopting the Parties' five page agreement on page limits, the Court declines to incorporate the text directly into TI's motion.

[2] TI repeatedly discusses the "MSP430" and "MSC121" identifiers in its motion, but Keranos does not dispute that these identifiers are excluded under the Court's orders.

For the reasons assigned by the Court in the Court's Prior Orders, the Court hereby finds that the identifiers "CC2430," "CC2510Fx," and "CC2511Fx" represent entire product families and are not properly definite under the Court's Prior Orders.

The Court accepts the Parties' agreement as to the TMS470RX1 and TIMSC1211Y4 products.

SIGNED this 3rd day of April, 2014.

ROY S. PAYNE
UNITED STATES MAGISTRATE JUDGE

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

KERANOS, LLC, Plaintiff, v. ANALOG DEVICES, INC., et al., Defendants.	Case No. 2:13-cv-00018-MHS-RSP

ORDER

THE COURT, having denied plaintiff Keranos' Motion for Leave to Amend Infringement Contentions and Keranos' Motion to Reconsider [Order, Dkt. 56], and having concluded in its Memorandum and Order dated April 3, 2014, that certain products of Defendants Apple Inc. ("Apple"), International Business Machines Corporation ("IBM") and Texas Instruments Incorporated ("TI") are excluded [Memorandum and Order, Dkt. 78];

FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against the "945 Express chipset" and "82573E LAN" products of Defendant Intel Corporation ("Intel");

AND FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against the "8 GB 2nd Generation iPod Nano" products of Defendant Apple;

AND FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against the ADUC814, ADUC824, and ADUC831 products of Defendant Analog Devices, Inc. ("ADI");

AND FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against the TMS470RX1 and TIMSC1211Y4 products of Defendant TI;

AND FURTHER, Keranos having agreed to dismissal with prejudice of its patent infringement claims against the LPC2888, LPC2000, LPC2917/2919/01, and PNX0161 products of Defendant NXP Semiconductor USA, Inc. ("NXP");

AND FURTHER, as a result of the above Orders and dismissals, Keranos is prohibited from proceeding on any patent infringement claims under U.S. Patent Nos. 4,868,629, 4,795,719, and 5,042,009 ("the Patents-in-Suit") against Defendants Apple, ADI, Intel, IBM, TI, NXP, and National Semiconductor Inc. ("NSC") (collectively, "Defendants");

AND FURTHER, Defendants having agreed to dismiss without prejudice their declaratory judgment claims and counterclaims of non-infringement and invalidity of the Patents-in-Suit;

IT IS HEREBY ORDERED that Keranos' patent infringement claims against Defendant Intel's "945 Express chipset" and "82573E LAN" products are dismissed with prejudice.

IT IS HEREBY ORDERED that Keranos' patent infringement claims against Defendant TI's TMS470RX1 and TIMSC1211Y4 products are dismissed with prejudice.

IT IS HEREBY ORDERED that Keranos' patent infringement claims against Defendant NXP's LPC2888, LPC2000, LPC2917/2919/01, and PNX0161 products are dismissed with prejudice.

IT IS HEREBY ORDERED that Keranos's patent infringement claims against Defendant Apple's "8 GB 2nd Generation iPod Nano" products are dismissed with prejudice.

IT IS HEREBY ORDERED that Keranos's patent infringement claims against Defendant ADI's ADUC814, ADUC824, and ADUC831 products are dismissed with prejudice.

IT IS FURTHER ORDERED that summary judgment against all remaining claims of infringement of the Patents-in-Suit against Defendants is granted for the reasons previously assigned, and these claims are dismissed with prejudice subject to Keranos' right to appeal.

IT IS FURTHER ORDERED that Defendants' declaratory judgment claims and counterclaims of non-infringement and invalidity of the Patents-in-Suit are dismissed without prejudice.

IT IS SO ORDERED.

SIGNED this 1st day of May, 2014.

MICHAEL H. SCHNEIDER
UNITED STATES DISTRICT JUDGE

United States Patent [19]
Eitan

[11] **Patent Number: 4,795,719**
[45] **Date of Patent: Jan. 3, 1989**

[54] **SELF-ALIGNED SPLIT GATE EPROM PROCESS**

[75] Inventor: **Boaz Eitan**, Sunnyvale, Calif.

[73] Assignee: **WaferScale Integration, Inc.**, Fremont, Calif.

[21] Appl. No.: **900,065**

[22] Filed: **Aug. 22, 1986**

Related U.S. Application Data

[62] Division of Ser. No. 610,369, May 15, 1984, Pat. No. 4,639,893.

[51] **Int. Cl.**[4] **H01L 27/10**
[52] **U.S. Cl.** **437/43**; 437/44; 437/48; 437/49; 437/52; 437/200; 437/984
[58] **Field of Search** ..... 357/23.5, 54, 71; 148/DIG. 141, DIG. 109; 29/576 B, 571, 577 C, 578

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,122,544	10/1978	McElroy .	
4,142,926	3/1979	Morgan	357/23.5
4,173,318	11/1979	Bassous et al.	29/571
4,173,791	11/1979	Bell	357/23.5
4,247,558	5/1981	Guterman	357/23.5
4,257,832	3/1981	Schwabe et al.	357/23.5
4,274,012	6/1981	Simko .	
4,297,719	10/1981	Hsu	357/23.5
4,318,216	3/1982	Hsu	437/29
4,334,292	6/1982	Eitan	357/23.5
4,336,603	6/1982	Kotecha et al.	357/23.5
4,380,866	4/1983	Countryman et al.	437/48
4,412,311	10/1983	Miccoli et al.	357/23.5
4,426,764	1/1984	Kosa et al.	29/571
4,462,090	7/1984	Iizuka .	
4,471,373	9/1984	Shimizu et al.	357/23.5
4,495,693	1/1985	Iwahashi et al.	29/576 B
4,561,004	12/1985	Kuo .	

FOREIGN PATENT DOCUMENTS

816931	7/1969	Canada .	
0045578	2/1982	European Pat. Off. .	
0158078	12/1982	Fed. Rep. of Germany	29/571
2437676	9/1979	France .	
0063684	5/1977	Japan	357/23.5
0089686	8/1978	Japan	357/23.5
0156369	12/1980	Japan	357/23.5
0071971	6/1981	Japan	29/571
0076878	5/1982	Japan	357/23.5
0206165	12/1983	Japan	357/23.5
2073484	10/1981	United Kingdom	357/23.5

OTHER PUBLICATIONS

IEEE Transactions, Electron Devices, vol. ED-29, No. 4, Apr. 1982, pp. 611-et seq., "Semiconductor MOSFET Structure for Minimizing Hot Carrier Generation".
IEEE Transactions, Electron Devices/vol. ED-32, No. 3, Mar. 1985, pp. 562-et seq., "Optimum Design of N+-N-Double-Diffused Drain MOSFET to Reduce Hot-Carrier Emission".
Electronics/Aug. 21, 1986, pp. 53-56.
Electronics/Sep. 4, 1986, p. 30.
Electronics/ Mar. 3, 1988, pp. 47-48.
IEE Transactions, Electron Devices/vol. ED-32, No. 5, May 1985, pp. 896-et seq., "Lightly Doped Drain Transistors for Advanced VLSI Circuits".

Primary Examiner—Brian E. Hearn
Assistant Examiner—Tom Thomas
Attorney, Agent, or Firm—Alan H. MacPherson; Gideon Gimlan; Forrest E. Gunnison

[57] **ABSTRACT**

A self-aligned split gate single transistor memory cell structure is formed by a process which self aligns the drain region to one edge of a floating gate. The portion of the channel underneath the floating gate is accurately defined by using one edge of the floating gate to align the drain region. The control gate formed over the floating gate controls the portion of the channel region between the floating gate and the source to provide split gate operation. The source region is formed sufficiently far from the floating gate so that the channel length between the source region and the closest edge of the floating gate is controlled by the control gate but does not have to be accurately defined.

8 Claims, 4 Drawing Sheets

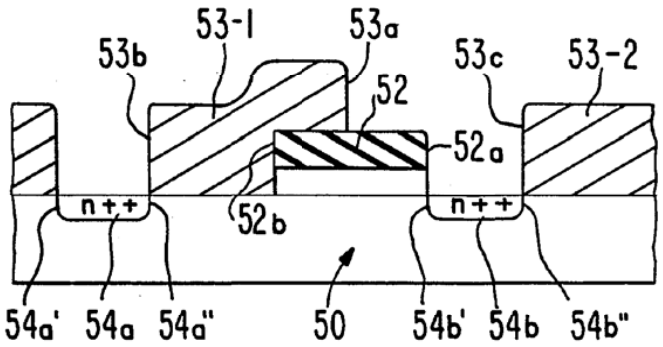


FIG. 1
PRIOR ART

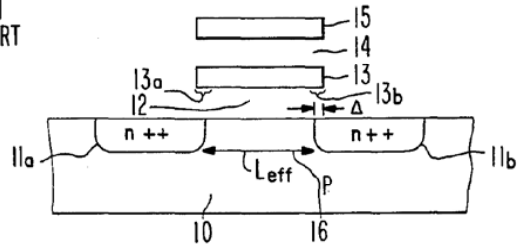


FIG. 2
PRIOR ART

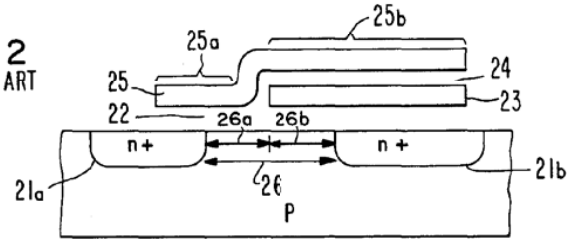


FIG. 3

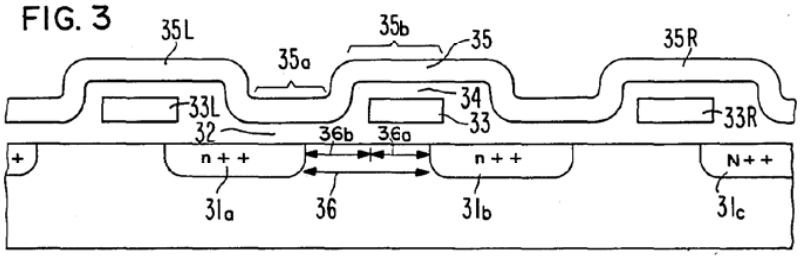
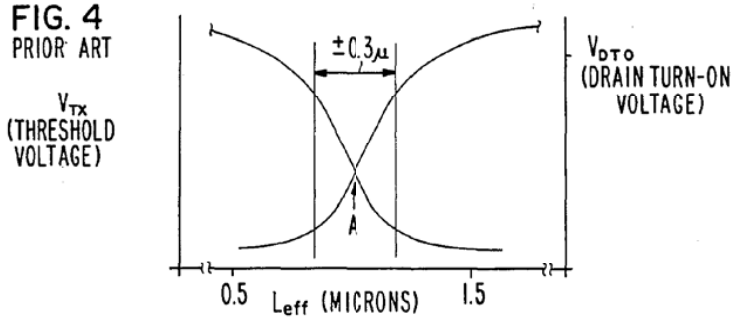


FIG. 4
PRIOR ART



KRNS000890

FIG. 5a

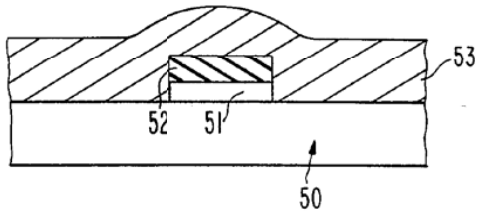
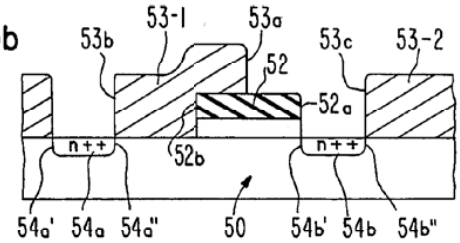


FIG. 5b



4,795,719

FIG. 6a

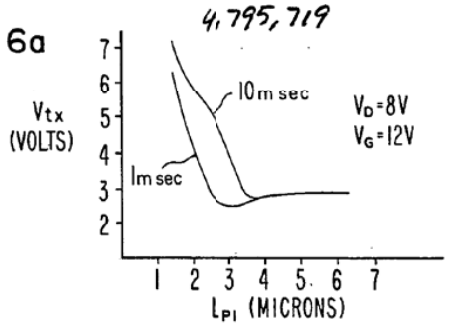
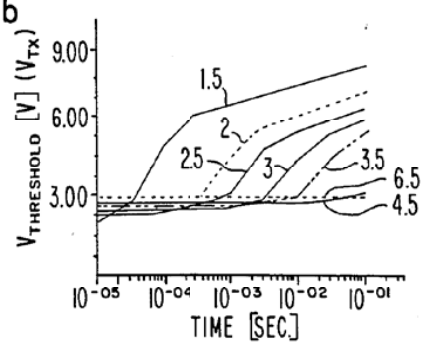


FIG. 6b



KRNS000891

FIG. 6c

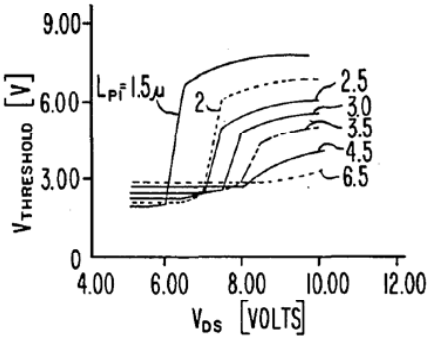
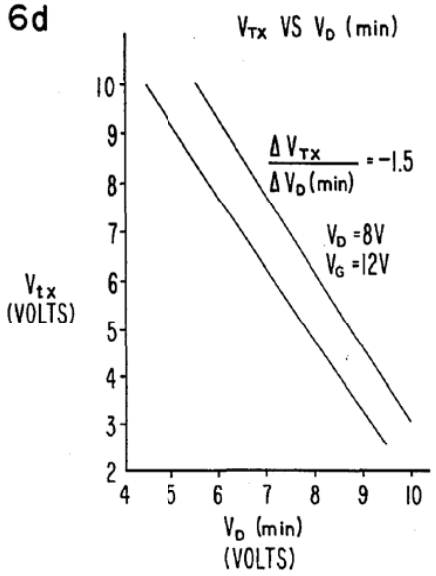


FIG. 6d



KRNS000892

FIG. 7a

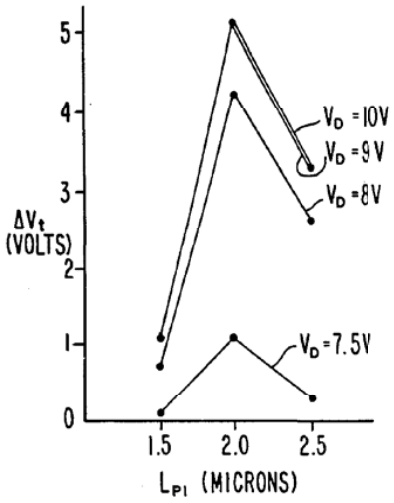


FIG. 7b

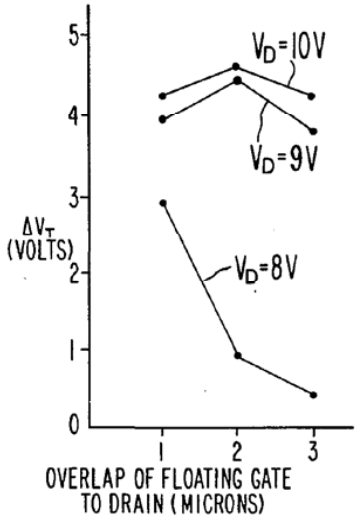
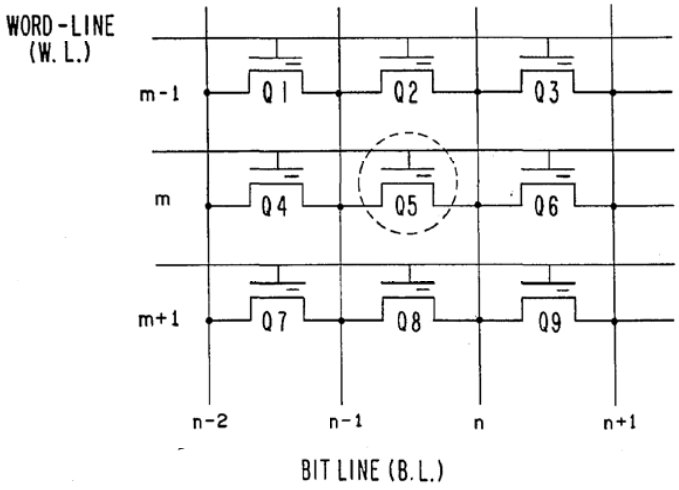


FIG. 8



KRNS000893

SELF-ALIGNED SPLIT GATE EPROM PROCESS

This application is a division of application Ser. No. 610,369 filed May 15, 1984, now U.S. Pat. No. 4,639,893.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a nonvolatile EPROM and more particularly to such an EPROM having a split gate (i.e., both a floating gate and a control gate) for controlling the writing and reading of each cell wherein the floating gate is self-aligned with the drain and the channel underlying the floating gate and the control gate is not self-aligned.

2. Prior Art

A split gate nonvolatile EPROM with increased efficiency is disclosed in U.S. Pat. No. 4,328,565 issued May 4, 1982 on an application of Harari, filed Apr. 7, 1980. As disclosed by Harari, the floating gate in an n channel EPOM cell extends over the drain diffusion and over a portion of the channel thereby to form a "drain" capacitance between the drain and the floating gate and a "channel" capacitance between the channel and the floating gate. A control gate then overlaps the floating gate and extends over the remainder of the channel near the source diffusion thereby to form a "control" capacitance between the floating gate and the control gate. These three capacitances form the coupling for driving each cell. The inversion region in the channel directly under the control gate is established directly by a "write or read access" voltage applied to the control gate. The inversion region in the channel directly under the floating gate is established indirectly through the drain and control capacitances and the channel capacitance by the control gate voltage and by another write access voltage applied to the drain. A cell is erased either by ultraviolet illumination or by electrons from the floating gate tunneling through a region of thinned oxide. The nonsymmetrical arrangement of the control gate and floating gate with respect to source and drain allows a very dense array implementation. Other split gate structures are disclosed in an article by Barnes, et al. entitled "Operation and Characterization of N-Channel EPROM Cells", published in Solid State Electronics, Vol. 21, pages 521–529 (1978) and an article by Guterman, et al. entitled "An Electrically Alterable Nonvolatile Memory Cell Using a Floating-Gate Structure", published in the IEEE Journal of Solid-State Circuits, Vol. SC-14, No. 2, April 1979.

FIG. 1 illustrates a typical EPROM of the prior art. In FIG. 1 a memory cell comprises n++ source region **11***a* and n++ drain region **11***b* separated by channel region **16**. Channel region **16** has an effective length L_{eff} as shown. Overlying channel region **16** is gate dielectric **12** on which is formed a floating gate **13**. Typically floating gate **13** is formed of polycrystalline silicon. Overlying floating gate **13** is insulation **14**, typically thermally grown silicon dioxide. Control gate **15** is formed above floating gate **13** on insulation **14**. The state of the transistor in FIG. **1** is determined by the charge placed on floating gate **13**. When electrons are placed on floating gate **13**, the threshold voltage V_{th} required on gate **15** to turn on the transistor (i.e., to form an n channel between source **11***a* and drain **11***b* thereby allowing current to flow from one to the other) is much greater than when no electrons are placed on

floating gate **13**. As shown in FIG. **1**, regions **13***a* and **13***b* of floating gate **13** overlie the source **11***a* and drain **11***b*, respectively, by a small amount "Δ". Consequently, a capacitance is formed between the source **11***a* and floating gate region **13***a* and between the drain **11***b* and floating gate region **13***b*. If the overlap by gate **13** of the source **11***a* and the drain **11***b* is the amount "Δ", then the capacitance C_{pp} between the floating gate **13** and the control gate **15** (both made of polycrystalline silicon) is given by the following equation:

$$C_{pp} \alpha A_{pp} \alpha W(L_{eff} + 2\Delta_{FG,D}) \quad (1)$$

In equation 1, C_{pp} is the capacitance between the floating gate **13** and the overlying control gate **15** (this capacitance is proportional to A_{pp}) and A_{pp}, the area of the floating gate **13**, is just the width W of the floating gate **13** (perpendicular to the sheet of the drawing) times the length of the floating gate **13** which is $(L_{eff} + 2\alpha_{FG,D})$.

The capacitance C_{PROM} between the floating gate **13** and the substrate **10** is proportional to the effective width W_{eff} (i.e. the width perpendicular to the sheet of the paper of the active area underneath the floating gate **13**) of the floating gate **13** times L_{eff}. Thus the capacitance C_{PROM} is

$$C_{PROM} \alpha A_{PROM} \alpha W_{eff}(L_{eff}) \quad (2)$$

The capacitive coupling $C_{FG,D}$ of the floating gate **13** to the drain **11***b* is given by

$$C_{FG,D} \alpha A_{FG,D} \alpha W_{eff}(\Delta_{FG,d}) \quad (3)$$

The coupling ratio $CR_{FG,D}$ of the capacitive coupling $C_{FG,D}$ of the floating gate **13** to drain **11***b* to the capacitive coupling C_{pp} of the floating gate **13** to the control gate **15** and the capacitive coupling C_{PROM} of the floating gate **13** to the substrate **10** is

$$CR_{FG,D} \alpha W_{eff}(\Delta_{FG,D}) / [W_{eff}(L_{eff}) + W\text{-}(L_{eff} + 2\Delta_{FG,D})] \quad (4)$$

As L_{eff} becomes smaller and smaller the impact of the coupling of the drain on the performance of the PROM cell becomes greater and greater until in the limit, as L_{eff} becomes very, very small, this coupling approaches 0.3 (taking into account different oxide thicknesses and the difference between W and W_{eff}, for example). The overlay "Δ" depends on the process and is substantially fixed.

FIG. **2** shows the prior art split gate structure as illustrated by Harari in U.S. Pat. No. 4,328,565 issued May 4, 1982. The major concern in this structure relates to the length of portion **26***b* of channel **26** beneath floating gate **23**. The structure of FIG. **2** is a nonself-aligned split gate structure. The total effective channel length **26** is defined by one mask and therefore is constant. Unfortunately, the length of the portion **26***b* of channel **26** beneath the floating gate **23** varies with mask alignment tolerances. Thus the effective channel length **26***b* depends strongly on the alignment process. As a result the best technology available today yields an effective tolerance of channel length **26***b* no better than ±0.5 to ±0.6 microns. For a typical nominal one micron effective channel length **26***b* the actual channel length will vary, due to manufacturing tolerances, over the range of about 1±0.6 micron. The result is a very wide variation in performance from one transistor memory cell to

the next. Programming and read current are both very sensitive to channel length. Good cells will be perfect but bad cells will not work. A good device has an effective channel **26***b* (in one embodiment 0.8 microns) which lies between a too-short channel length (for example 0.2 microns or less, so that considering manufacturing variations there may not be any overlap at all of floating gate **23** over the channel **26** and thus there will be no programming of the cell) and a too-long channel length (for example greater than 1.4. microns) with unacceptably slow programming. The major issue in this prior art structure thus is the length of channel portion **26***b* (L_{eff}) rather than the coupling. Therefore in a structure such as that shown in FIG. **2** there can be coupling between drain **21***b* and floating gate **23** but if the channel length **26***b* is not carefully controlled, the memory cell is not going to perform as expected.

A major problem in the prior art EPROM of FIG. **1** relates to the relationship between the program threshold voltage V_{tx} and the drain turn on voltage V_{DTO} of the device. V_{DTO} is the voltage on the drain which, when capacitively coupled to the floating gate **13**, turns on the transistor. As shown in FIG. **4**, for L_{eff} as shown in FIG. **1** increasing from about 0.5 to 1.2 microns, the program threshold V_{tx} drops below the acceptable program threshold. On the other hand the drain turn-on voltage V_{DTO} becomes as high as the junction breakdown voltage for L_{eff} greater than about one micron. Below one micron, V_{DTO} is very low and may go as low as three to five volts which causes the array of EPROMS to fail. The crossover point is shown as "A" in FIG. **4**. In designing a regular EPROM, the crossover point A should be such that V_{tx} is high enough (i.e. greater than five volts) while V_{DTO} is not too low (i.e. not lower than eight volts). However, both curves V_{DTO} and V_{tx} are quite steep at the crossover point A and thus the characteristics of the device are very sensitive to L_{eff}. So if the tolerance on L_{eff} is even ± 0.3 microns, which is very good, then the characteristics of the device are still relatively unpredictable. Obviously the desired solution is to eliminate the effect of V_{DTO} and optimize L_{eff} for V_{tx}.

SUMMARY OF THE INVENTION

In accordance with my invention, I overcome the problems of the prior art by providing a memory cell using a split gate structure containing both a control gate and a floating gate in which the floating gate is self-aligned to the drain region. The control gate is not self-aligned. By "self-aligned" I mean that the portion of the transistor channel length under the floating gate will be defined by the floating gate itself regardless of any processing misalignments thereby insuring a constant channel length under the floating gate. To do this, a special process is employed wherein the floating gate is used to define one edge of the drain region. The source region is defined at the same time as the drain region but the alignment of the source region relative to the floating gate is not critical so long as the source region does not underlie and is spaced from the floating gate.

In a process in accordance with this invention, the diffused drain region (which also functions as a bit line and which corresponds to an elongated drain region of the type shown in the above-mentioned '565 patent) is formed using the floating gate to define one edge of the drain region. In the preferred embodiment, the drain and source regions are formed by ion implantation and one edge of the floating gate defines the lateral limit of one side of the drain region. A photoresist material partially extends over the floating gate in one direction and beyond the floating gate in the other direction and the source region is defined by an opening in the portion of this photoresist extending beyond the floating gate in the other direction. The result is to form a precisely defined channel portion L_{eff} of the channel region beneath the floating gate and a remaining relatively imprecisely defined portion of the channel region (to be controlled by a to-be-formed control gate electrode which is part of the word line) underneath the photoresist between the other edge of the floating gate and the source region.

In accordance with my invention, any misalignment between the floating gate and the source region is covered by a to-be-formed control gate and has little effect on the operation of the memory cell while the floating gate is self-aligned to the drain region.

This invention will be understood in more detail in conjunction with the following drawings:

DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a prior art EPROM using a single floating gate beneath the control gate;

FIG. **2** illustrates the split gate structure of the prior art wherein the floating gate is not self-aligned to the drain region and the control gate is formed over part of the channel region;

FIG. **3** illustrates the split gate structure of this invention wherein the floating gate is self-aligned to the drain region and overlies but is insulated from an accurately defined portion L_{eff} of the channel region between the source and drain and the control gate overlies the floating gate and that portion of the channel region not overlain by the floating gate but is insulated therefrom;

FIG. **4** illustrates the relationship between threshold voltage V_{tx} and drain turn-on voltage V_{DTO} for the structure of FIG. **1**;

FIGS. **5***a* and **5***b* illustrate the novel process which I use to manufacture the novel self-aligned split gate structure of my invention;

FIGS. **6***a* through **6***d* illustrate the effect of the channel length L_{P1} under the floating gate on programming;

FIGS. **7***a* and **7***b* show the tight envelope of operation for the nonself-aligned structure and illustrate graphically the advantages of my self-aligned split gate structure; and

FIG. **8** shows in schematic form a memory array formed using the self-aligned split gate structure of my invention.

DETAILED DESCRIPTION

The following detailed description is meant to be illustrative only and not limiting. Other embodiments of this invention will be obvious to those skilled in the art in view of the following description. In FIGS. **5***a* and **5***b*, only the cross-section of a single memory cell or a portion thereof is shown while in FIG. **3** two cells and part of a third are shown in cross-section. It should be understood that a semiconductor integrated circuit memory made in accordance with this invention employs a plurality of such cells together with peripheral circuits for writing data into memory and for accessing the data stored in the memory. For simplicity these circuits are not shown.

The starting point for the process of my invention to yield my novel self-aligned split gate structure is the

same as in the nonself-aligned split gate structure of the prior art and in particular of the '565 patent. Thus as shown in FIG. **5***a* a silicon substrate **50** typically having a resistivity of 10–50 ohm-centimeters has formed thereon in a standard manner a layer of gate oxide **51**. Gate oxide **51**, typically 300 angstroms thick, then has formed on it a first layer of polycrystalline silicon (often called "poly 1") which is patterned as shown in FIG. **5***a* to form a floating gate **52**. The oxide **51** beneath the portions of polycrystalline silicon removed to form floating gate **52** is then removed by an etching process (typically a plasma etch) and a photoresist layer **53** is then formed over the top surface of the structure.

As shown in FIG. **5***b* photoresist layer **53** is then patterned so that a particular segment **53-1** of photoresist is formed to partially overlie floating gate **52**. Photoresist **53-1** has a right edge **53***a* which is formed to overlie the floating gate **52** somewhere near its middle and a left edge **53***b* which is formed to the left of left edge **52***b* of floating gate **52**. The width of floating gate **52** is typically 1.5 to 2 microns and thus it is not difficult to ensure with sufficient certainty given typical tolerances in the manufacturing process that edge **53***a* is to the left of right edge **52***a* of floating gate **52** even for a reasonably expected worst case mask misalignment during the manufacturing process. It is also quite simple to insure that left edge **53***b* is sufficiently to the left of left edge **52***b* of floating gate **52** so that left edge **52***b* of floating gate **52** is never exposed, even in a worst case alignment mismatch of masks during manufacture. Thus the to-be-formed source **54***a* will always be laterally spaced from the left edge **52***b* of floating gate **52**.

Following the formation of patterned photoresist **53-1** the structure is subjected to an ion implantation at a selected well-known dosage (typically 4×10^{12} per cm^2) to form n++ drain region **54***b* and n++ source region **54***a* in the top surface of the semiconductor material **50**. The region **54***b* has its left edge **54***b'* defined by the right edge **52***a* of floating gate **52** and its right edge **54***b''* defined by the left edge **53***c* of patterned photoresist **53-2**. The source region **54***a* has its right edge **54**a'' defined by the left edge **54***b* of patterned photoresist **53-1**. Thus the drain region **54***b* is self-aligned to the right edge **52***a* of floating gate **52**. However, the right edge **54***a''* of source region **54***a* is self-aligned to the left edge **53***b* of photoresist **53-1**. The uncertainty in the location of left edge **53***b* of patterned photoresist **53-1** relative to left edge **52***b* of floating gate **52** represents an uncertainty in the length of the control gate channel (corresponding to channel portion **36***b* in FIG. **3**) and not of the floating gate channel L_{eff} (corresponding to channel portion **36***a* in the center cell of FIG. **3**). By placing a proper voltage on the to-be-formed control gate, (corresponding to control gate **35** in FIG. **3**), the channel length under the control gate becomes irrelevant and the conduction or nonconduction of the total channel is determined by the voltage placed on the floating gate **52** (corresponding to floating gate **33** in FIG. **3**). Because floating gate **52** is uniformly coupled to drain **54***b* in all transistors in a memory array made in accordance with this invention and further because the effective channel length L_{eff} (corresponding to channel **36***a* in FIG. **3**) is substantially the same underneath all floating gates **52** in all transistors in a memory array formed in accordance with this invention, the structure of this invention yields a split gate programmable EPROM capable of being manufactured with much higher yield than the prior art EPROMs.

The remaining steps in the process are standard well known steps in the silicon gate EPROM technology. Insulation (not shown) is formed over floating gate **52**. A control gate (often called "poly 2") corresponding to control gate **35** in FIG. **3** is formed, usually as part of a word line. The resulting structure appears as shown in FIG. **3**. FIG. **3** shows floating gate **33** and floating gates **33L** and **33R** formed to the left and right of floating gate **33**. All three floating gates have their right edges self-aligned to the left edges of the underlying drain regions in accordance with this invention. The drain region for a given cell doubles as the source region for the cell to the right.

The finished structure made by the process of this invention as illustrated in FIGS. **5***a* and **5***b* is shown in FIG. **3**. In FIG. **3** the floating gate **33** has been formed prior to the formation of the source and drain regions **31***a* and **31***b*. The floating gate **33** is formed on a thin layer of insulation overlying a portion of the to-be-formed channel region between the source and drain. The right edge of the floating gate **33** has been used to define one edge of the drain region **31***b*. Overlying floating gate **33** is insulation **34** (typically silicon oxide) and overlying oxide **34** is the control gate **35**. A portion **35***a* of control gate **35** overlies a second portion of the channel region between the left end of the floating gate **33** and the source region **31***a*. As described herein, the channel region **36***b* beneath portion **35***a* of control gate **35** can have a length **36***b* which varies substantially without affecting the performance of the device.

The structure shown in the central part of the cross-section in FIG. **3** is but one cell of a plurality of such cells. In a typical virtual ground structure the drain **36***b* for the cell shown in cross-section in FIG. **3** serves as the source for another cell located just to the right. Likewise the source **36***b* serves as the drain for a second cell located just to the left. The portions of the floating gates **33**L and **33**R associated with these adjacent cells are shown in FIG. **3**.

Note that the floating gate **52** (FIG. **5***a* and **5***b*) becomes capacitively coupled to drain region **54***b* by the lateral diffusion of left edge **54***b'* beneath floating gate **52** during further processing of the structure. This lateral diffusion is typically around 0.3 microns. However contrary to the prior art, the floating gate **52** is formed before the formation of the drain region **54***b*, rather than after, and is precisely self-aligned to one edge of the drain region **54***b*.

FIG. **6***a* illustrates the variation in threshold voltage versus the drain channel length (L_{P1}) of the floating gate ("poly 1"). In FIG. **6***a* the ordinate is the program threshold and the abscissa is the length of the floating gate channel L_{P1} in microns. (Of importance, FIGS. **6***a* through **6***d* and **7***a* and **7***b* use drawn dimensions. However, the channel lengths **36***a* and **36***b* shown in FIG. **3** are the effective dimensions after processing. Thus channel length **36***a* is denoted L_{eff} to represent the effective length of this channel after processing, while before processing this channel length is a drawn dimension and as such is denoted by the symbol L_{P1}. Accordingly, each of the dimensions L_{P1} shown in FIGS. **6***a* through **6***d* and **7***a* and **7***b* must be corrected (i.e., reduced) by a given amount (approximately 0.5 microns), to reflect the effect of processing. Naturally the amount of the correction will vary with the processing.) The threshold voltage V_{tx} obtained or programmed in a given time for a given drain voltage and gate voltage (corresponding in FIG. **6***a* to a drain voltage of 8 volts and a gate

voltage of 12 volts) drops rapidly as the length of the channel L_{P1} under the floating gate **52** (FIG. **5***b*) increases to a minimum V_{tx} of about 2.5 volts for L_{P1} of somewhere between 3 to 4 microns and then increases slightly. This minimum V_{tx} corresponds to the initial device threshold before programming. The threshold V_{tx} represents the voltage which must be applied to the control gate (such as gate **35** in FIG. **3**) to turn on the transistor beneath the control gate as shown in FIG. **3** when the cell containing that transistor has been programmed. Thus as the length of the channel **36***a* underneath the floating gate **33** increases (FIG. **3**) the threshold voltage necessary to turn on the transistor and create a channel from the source region **31***b* to the drain **31***a* decreases. As is shown in FIG. **6***a*, both 1 millisecond and 10 millisecond programming times yield substantially the same shaped curve.

FIG. **6***b* illustrates the effect of the length of the channel **36***a* underneath floating gate **33** on the threshold voltage (ordinate) versus programming time (abscissa). The various curves reflect different lengths L_{P1} of the channel **36***a* (FIG. **3**) beneath floating gate **33** in microns. As these channel lengths increase, the threshold voltage for a given programming time drops. Thus for a programming time of 10^{-2} seconds, the threshold voltage for a 1.5 micron channel length L_{P1} is approximately 7 volts whereas the threshold voltage for a 3.0 micron channel $L_{P1}$is about 4 volts. These curves were obtained for a voltage V_{DS} from the drain to the source of 8 volts and a voltage on the control gate **35** of 12 volts. The curves of FIG. **6***b* illustrate that the shorter the floating gate the stronger the field which is formed and therefore the greater the number of electrons which are placed on the floating gate thereby resulting in a larger threshold voltage V_{tx} to turn on the transistor.

FIG. **6***c* is a plot of threshold voltage V_{tx} (ordinate) versus the voltage on the drain **31***b* (FIG. **3**) with the length of channel **36***a* beneath floating gate **33** as the parameter on the various curves. For a given drain voltage V_D (for example 8 volts) the threshold voltage V_{tx} goes up as the length L_{P1} of the channel **36***a* beneath floating gate **33** goes down. The curves of FIG. **6***c* were taken with a control channel L_{P2} (corresponding to the drawn dimension of channel **36***b* in FIG. **3**) beneath the control gate **35** of 2.5 microns, a gate voltage on control gate **35** of 12 volts and a programming time of 10 milliseconds (10^{-2} seconds). These curves illustrate that once a given drain voltage difference V_{DS} is achieved between the drain and the source, increasing the drain voltage beyond a given amount has substantially little effect on the threshold voltage V_{tx} of the transistor. In other words, $\Delta V_{tx}/\Delta V_{DS}$ becomes substantially zero thereby showing that increasing the drain voltage coupled to the floating gate has little effect on the programming of the transistor. Thus after the program threshold voltage V_{tx} is reached, increasing the drain to source voltage V_{DS} does not achieve any significant improvement in performance.

As L_{P1} increases, the threshold voltage V_{tx} at which ΔV_{tx} over ΔV_{DS} becomes very small decreases. So increasing V_{DS} does even less for structures with longer floating gates.

In FIG. **6***c* each consecutive point on a given line for a given L_{P1} represents an additional 10 milliseconds of programming time rather than just 10 milliseconds of programming time. Accordingly the curves for V_{tx} versus V_{DS} in FIG. **6***c* would be even flatter than shown in FIG. **6***c* if a constant programming time was applied to program the cell from different V_{DS} start points.

FIG. **6***d* illustrates the very tight predictability of threshold voltage V_{tx} versus V_D (min) for the structure of this invention. V_D (min) is defined as the minimum V_{DS} needed to start programming (i.e., to start efficient electron flow onto the floating gate). In FIG. **6***c* V_D(min) is the V_{DS} at which the curve shows a break point sharply to the right. This break point or "knee" corresponds to the V_D(min) plotted in FIG. **6***d*.

The relationship of FIG. **6***d* to FIG. **6***c* illustrates a basic point of my invention. In a 256K EPROM the time to program the cells in the-EPROM theoretically equals 256K times the time to program each cell divided by 8 (ROMs are programmed one byte at a time). Therefore, if the programming time of each cell can be significantly reduced, the efficiency of programming a large number of EPROMs can be proportionally increased. I have discovered that to program to a given threshold voltage V_{tx} in a given programming time, the key is to control the length of L_{P1} and in particular to make this length (which is related to the channel **36***a* in FIG. **3**) as small as practical without generating punch-through from the source to the drain. As shown by analysis of FIG. **6***d*, the threshold voltage V_{tx} is increased for a given programming time by decreasing V_D(min). As shown in FIG. **6***c* V_D(min) decreases as L_{P1} decreases in length. Accordingly, decreasing L_{P1} is the key to programming to a given threshold voltage V_{tx} in a given time. My invention not only allows a small effective channel length L_{eff} to be achieved beneath the floating gate but allows this channel length to be achieved in a controllable and reproducible manner throughout an EPROM array thereby to obtain repeatable and consistent results throughout the array.

FIG. **7***a* illustrates change of threshold voltage, ΔV_T for three different L_{P1} (i.e., three different drawn channel lengths beneath the floating gate) for the structure shown in FIG. **2**. In a nonself-aligned structure, the proper length of the channel under the floating gate is crucial to achieve maximum threshold voltage V_{tx}. As shown in FIG. **7***a* if the channel length **36***a* becomes too short (for example, 1.5 microns), then punch-through occurs between the source **31***a* and drain **31***b* during programming resulting in a failure to program the device. The proper alignment of a floating gate in the nonselfaligned structure to optimize the length of the channel **36***a* beneath the floating gate **33** and the overlap of the floating gate to drain is crucial. The very sharp peak in FIG. **7***a* reflects the variation in V_{tx} with channel length L_{P1}. FIG. **7***a* shows that to optimize the device for the minimum channel length L_{P1} in terms of programming efficiency results in a lower initial threshold before programming and higher final threshold after programming so as to obtain a higher read current. This means a lower impedance in the circuit which in turn means that during read a capacitor in the sense amplifier in the peripheral circuitry of the memory discharges faster through a programmed transistor than otherwise would be the case resulting in shorter access time.

Three effective channels beneath the floating gate (1.5 micron, 2.0 micron and 2.5 micron) are shown in FIG. **7***a*. The parameter ΔV_T (representing the change in threshold voltage as a function of different channel length) is illustrated by the curves. This change in voltage is particularly pronounced as one goes from 1.5 to 2 to 2.5 micron length for L_{P1}. The change in V_{tx} as a function of channel length is similar to that shown in

FIG. 6a for the self-aligned structure of my invention. However, as one goes from a 2 micron L_{P1} to 1.5 micron L_{P1} and shorter, a new phenomenon appears reflecting possible punch through from the source to the drain and V_{tx} thus is lower than would be expected. The nonselfaligned curve shows that a proper L_{P1} is critical to obtaining a predicted threshold voltage. However, with nonself-aligned floating gate technology L_{P1} can vary even across a given chip causing a variation in V_{tx} from cell to cell within a given memory. Often this variation is unacceptable. As can be seen by the curves of FIG. 7a, a given memory can have L_{P1} from cell to cell varying for example from 1.5 microns all the way to 2.5 microns or greater because of misalignment in the masking during the processing of the wafer. Accordingly, V_{tx} is unpredictably variable across the wafer often resulting in unacceptable performance.

FIG. 7b shows the effect of overlap and V_D on threshold voltage. For the nonself-aligned device the structure must be aligned so that the 3 sigma worst case of alignment gives a satisfactory channel length 36a beneath floating gate 33. Increasing the coupling between the floating gate and the drain does not improve the threshold voltage of the device for given programming conditions so overlapping the drain with the floating gate does not help. The more overlap of the floating gate to the drain means the more electrons required to charge the floating gate for a given channel length 36a beneath the floating gate. So instead of improving the efficiency of the device, increasing the overlap of the floating gate to the drain actually decreases this efficiency. A minimum overlap of the floating gate to the drain is needed to insure that accelerated electrons hit and lodge in the floating gate rather than in the control gate or the word line.

FIG. 7b shows that as the overlap of the nonselfaligned structure increases, the ΔV_T actually declines for a given V_D. Again, this shows that the coupling between the drain and the floating gate is not helpful to achieving a desired V_{tx} and indeed can even be harmful.

The circuit of this invention is highly scaleable and retains its self-aligned character as it is scaled.

An important effect of this invention is that by choosing the correct L_{P1} the programming time for a memory array can be substantially reduced. For example, a prior art 256K EPROM takes approximately 150 seconds or 2½ minutes to program. A 256K EPROM using the structure of this invention can be programmed in approximately 30 seconds. This is a substantial improvement resulting in lower programming costs and lower test costs.

An additional advantage flowing from this invention is that the uncertainty in the location of the floating gate due to mask alignment tolerances is substantially reduced compared to the uncertainty in the location of the floating gate in the prior art nonself-aligned structure and in the standard prior art EPROM (nonsplit gate with self-aligned). Table 1 illustrates this improvement with respect to the self-aligned split gate structure of this invention compared to the standard non-split gate self-aligned structure of the prior art.

TABLE I

	Standard EPROM (Non-split but self-aligned gate)	Self-aligned split gate structure of this invention
STEP 1	Poly 1 (Floating gate) Critical dimension not defined but non-critical dimensions are defined	1 (Floating gate) Critical dimension defined
STEP 2	Poly 2 (Control gate) Define critical dimensions of control gate - Accuracy degraded because of rough, nonplanar topology associated with two layers of polycrystaline silicon	
STEP 3	Poly 1 critical dimension defined using Poly 2 as a mask	

Table I compares only the critical steps in the two processes used to define the floating gate and thus the crucial channel length L_{eff}. L_{eff} is the important channel length in the self-aligned split gate structure of this invention and in any EPROM structure. Note that in a standard non-split gate self-aligned structure L_{eff} is the total channel length between the source and drain.

As shown in Table I three steps are required to define the critical dimension of the floating gate in the standard non-split gate self-aligned structure. In the first step only the noncritical dimensions corresponding to the width (but not the length) of the channel beneath the floating gate are defined. The critical dimensions of the floating gate corresponding to the channel length beneath the floating gate are not defined. In step 2 the second layer polycrystalline silicon from which the control gate will be fabricated is deposited. The critical dimension of this second layer (known as "poly 2") is defined in step 2. This dimension corresponds to the channel length between the to-be-formed source and drain regions. However, the accuracy with which the critical dimension of the control gate is fabricated is degraded because of the rough nonplanar topology associated with the two layers of polycrystalline silicon deposited on the wafer. In the third step the first layer of polycrystalline silicon (poly 1) has its critical dimension (corresponding to channel length L_{P1}) defined using the second layer of polycrystalline silicon as a mask. Again, the accuracy with which the critical dimension of the first layer of polycrystalline silicon is defined is degraded due to the uneven topology of the structure.

In contrast, the self-aligned split gate structure of my invention defines the critical dimension of the poly 1 floating gate layer in step 1.

As the above comparison shows, the channel length L_{P1} for the standard nonsplit gate self-aligned structure is equal to the drawn length of the channel plus or minus the uncertainty in the critical dimension associated with the poly 2 definition step plus or minus the uncertainty introduced in the critical dimension of the channel length associated with poly 1 using poly 2 as a mask. Thus the uncertainty in the effective channel length in the standard nonsplit gate self-aligned structure has two components introduced by two critical dimensions. On the other hand, using the self-aligned split gate structure of my invention, only one uncertainty in a critical dimension occurs and that occurs in the first step where the poly 1 critical dimension is defined and the topology is smooth. Accordingly my

invention yields a double processing advantage over the process by which the standard non-split gate self-aligned structure of the prior art is made by eliminating one critical dimension in defining L_{eff} and by introducing a much smoother topology during the formation of the critical channel length L_{eff}.

Table 2 compares the critical steps required to define the poly 1 floating gate in the nonself-aligned split gate structure of the prior art compared to the single step required to define the floating gate in the self-aligned split gate structure of my invention.

TABLE II

	Nonself-aligned split gate	Self-aligned split gate of this invention
STEP 1	Source and Drain Implanted	Poly 1 (Floating Gate) Define critical dimension
STEP 2	Poly 1 (Floating Gate) Define critical dimension	

Step 1 in fabricating the prior art nonself-aligned split gate structure is to implant the source and drain regions in the device. Step 2 is then to deposit the poly 1 layer and then form the floating gate from this layer. The critical dimension L_{P1} is defined by this step. Unfortunately, uncertainty in the length of LP_1 results from the uncertainty in the critical dimension of the poly 1 plus or minus the misalignment of the mask used to define the critical dimension of the floating gate relative to the underlying drain region. Typically the uncertainty in the critical dimension is ±0.3 microns while the uncertainty due to the mask misalignment is ±0.6 microns. When combined in a statistical sense (root mean square) the total uncertainty in L_{P1} can be ±0.6 or ±0.7 microns. To the contrary, using the self-aligned split gate structure of my invention, the critical dimension of the poly 1 floating gate is defined with an uncertainty at most of about ±0.3 microns. Accordingly, my invention achieves a substantial improvement in manufacturing accuracy over the prior art nonself-aligned split gate structure.

FIG. 8 illustrates an EPROM array fabricated using the self-aligned split gate structure of my invention. For simplicity, an array of nine (9) transistors or cells is shown. The programming and reading of cell or transistor Q5 will be described. Note that the array comprises of word line rows m-1, m and m+1 and bit line columns n-2, n-1, n and n+1. Column n-2 is the source of transistors Q1, Q4 and Q7 while column n-1 is the drain of transistors Q1, Q4, and Q7 and the source of transistors Q2, Q5 and Q8. Similarly, column n is the drain of transistors Q2, Q5 and Q8 and the source of transistors Q3, Q6 and Q9. Column n+1 is the drain of transistors Q3, Q6 and Q9.

In operation, to read device m,n (i.e. cell Q5) all bit lines except n-1 are set at 2 volts. Bit line n-1 is set at ground. Word line m is set at 5 volts while all other word lines except m are set at ground.

To program device m,n (i.e., cell Q5) all bit lines except n are set at ground while bit line n is set at 8 or 9 volts. All word lines except m are set at ground while word line m is set at 12 volts. During programming, device m,n+1 (i.e., cell Q6) is also in programming condition but in the reverse configuration (i.e., the high voltage is applied away from the floating gate). In this configuration there is no programming of m, n 1. This asymmetry in the split gate EPROM is what enables one to utilize the virtual ground approach.

While one embodiment of this invention has been described, other embodiments of this invention will be obvious to those skilled in the semiconductor arts in view of this disclosure.

What is claimed is:

1. A method of manufacturing a memory cell containing a split gate transistor comprising:

forming first polycrystalline silicon on, but separated from a semiconductor substrate by first insulation, said first polycrystalline silicon defining a floating gate having a first edge and a second edge opposite said first edge;

forming a photoresist pattern over said substrate and over a surface of said first polycrystalline silicon, said surface extending laterally between the first and second edges, a first opening being formed in said photoresist pattern to expose both the first edge of said floating gate and a first portion of the semiconductor substrate extending laterally from said first edge and a second opening being formed in said photoresist pattern to expose a second portion of the semiconductor substrate laterally spaced apart from said floating gate;

implanting selected impurities into those portions of the semiconductor substrate exposed by the openings of said photoresist thereby to form a source region laterally spaced apart from said floating gate and a drain region extending from but self-aligned to the first edge of said floating gate.

2. The method of claim 1 wherein said drain region has a selected edge self-aligned to the first edge of said floating gate.

3. The method of claim 1 wherein the first opening in the photoresist pattern is patterned to expose a laterally extending surface portion of the floating gate, said surface portion extending from the first edge to a point near the middle of the floating gate.

4. A method according to claim 1 further comprising:

forming second polycrystalline silicon to insulatively overlap the floating gate and a channel portion of the substrate located between a portion of the substrate overlapped by the floating gate and the second portion of the substrate, said second polycrystalline silicon defining a control gate of the split gate transistor.

5. A method for manufacturing a split gate transistor having an insulated floating gate overlying a channel region of the transistor and a control gate extending over the floating gate, the method comprising:

forming a first insulative layer extending laterally on a semiconductor substrate;

forming a first poly layer made of polycrystalline silicon on the first insulative layer;

patterning the first poly layer to include opposed first and second edges defining opposed ends of the floating gate;

forming a photoresist layer over the first poly layer;

patterning the photoresist layer so that the first edge of the first poly layer and a surface portion of the first poly layer extending laterally from the first edge are exposed to define a peripheral portion of a drain implantation window;

further patterning the photoresist layer to cover the second edge of the first poly layer and to extend laterally beyond the second edge of the first poly layer to terminate at an edge of the photoresist

layer to thereby define a peripheral portion of a source implantation window, the edge of the photoresist layer being positioned such that the source implantation window is spaced apart from the floating gate; and

implanting impurities through the source and drain implantation windows to form respective source and drain regions of the transistor, the drain region being self aligned thereby to the first edge of the floating gate and the source region being spaced apart from the floating gate.

6. A method according to claim 5 further comprising:

forming a second insulative layer extending laterally over the first poly layer; and

forming a second poly layer made of polycrystalline silicon on the second insulative layer, the second poly layer extending to insulatively overlay a channel portion of the transistor between the source and drain regions, said second poly layer defining a control gate of the split gate transistor.

7. A manufacturing method for assuring consistency over process variations in the effective channel length of a plurality of split gate transistors which are to be formed each to have a floating gate laterally spaced apart from a source region of the transistor, the method comprising:

insulatively disposing the floating gate of each transistor on a semiconductive substrate;

forming a photoresistive coating on the floating gate of each transistor, the coating extending laterally beyond the floating gate to cover the substrate;

creating a first opening in the coating to expose an edge portion of the floating gate of each transistor and a first portion of the substrate directly adjacent to the edge portion;

creating a second opening in the coating, laterally spaced apart from the floating gate, to expose a second portion of the substrate; and

implanting doping impurities through the first and second openings to create for each of the plurality of transistors a drain region which is self-aligned to the edge portion of the floating gate of the transistor and a source region which is spaced apart from the floating gate of the transistor.

8. A method according to claim 7 further comprising:

forming a control line to insulatively overlap the floating gates of each of the transistors and to further insulatively overlap channel portions of each of the transistors between the source and drain regions of the transistors, said control line defining a control gate for each of the transistors.

* * * * *

United States Patent [19]

Eitan

[11] **Patent Number: 4,868,629**

[45] **Date of Patent: Sep. 19, 1989**

[54] **SELF-ALIGNED SPLIT GATE EPROM**

[75] Inventor: **Boaz Eitan**, Sunnyvale, Calif.

[73] Assignee: **WaferScale Integration, Inc.**, Fremont, Calif.

[21] Appl. No.: **762,582**

[22] Filed: **Aug. 2, 1985**

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 610,369, May 15, 1984.

[51] **Int. Cl.**[4] **H01L 27/10**

[52] **U.S. Cl.** **357/45**; 357/41; 357/23.5; 357/23.9; 365/185

[58] **Field of Search** 357/23.5, 23.9, 41, 357/45; 365/185

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,122,544	10/1978	McElroy	357/23.5
4,142,926	3/1979	Morgan	357/23.5
4,173,791	11/1979	Bell	357/23.5
4,173,818	11/1979	Bassous et al.	357/23.5
4,257,832	3/1981	Schwabe et al.	357/23.5
4,267,558	5/1981	Guterman	357/23.5
4,274,012	6/1981	Simko	357/23.5
4,297,719	10/1981	Hsu .	
4,300,212	11/1981	Simko	357/23.5
4,318,216	3/1982	Hsu	357/23.5
4,328,565	5/1982	Harari	357/23.5
4,334,292	6/1982	Kotecha	357/23.5
4,336,603	6/1982	Kotecha et al.	357/23.5
4,412,311	10/1983	Miccoli et al.	357/23.5
4,426,764	1/1984	Kosa et al.	357/23.5
4,462,090	7/1984	Iizuka	365/185
4,471,373	9/1984	Shimizu et al.	357/23.5
4,495,693	1/1985	Iwahashi et al.	357/23.5
4,561,004	12/1985	Kuo et al.	357/41

FOREIGN PATENT DOCUMENTS

816931	7/1969	Canada	357/23.5
0045578	2/1982	European Pat. Off.	357/23.5
1647781	5/1985	European Pat. Off. .	
0158078	12/1982	Fed. Rep. of Germany	357/23.5
2437676	9/1979	France	357/23.5
0063684	4/1977	Japan	357/23.5
0089686	7/1978	Japan	357/23.5
54-156484	12/1979	Japan	357/23.5
0156369	12/1980	Japan	357/23.5
0071971	6/1981	Japan	357/23.5
0076878	5/1982	Japan	357/23.5
57-96572	6/1982	Japan	357/23.5
0206165	12/1983	Japan	357/23.5
2073484	10/1981	United Kingdom	357/23.5

OTHER PUBLICATIONS

Article entitled "High Density Flash EEPROMs Are About To Burst On The Market", pp. 47 and 48, *Electronics*, Mar, 3, 1988.

Shirota, Paper entitled "A New NAND Cell for Ultra High Density 5V only EPROM".

IEEE Transactions on Electron Devices, vol. ED-32, No. 5, 5/85 "Lightly Doped Drain Transistors For Advanced VLSI Circuits", pp. 896 et seq.

IEEE Trans. on EL.DV., vol. ED-29, No. 4, 4/82, By takeda et al, pp. 611 et seq., "Semiconductor MOSFET Structure For Minimizing Hot-Carrier Generation".

IEEE Trans. on EL. DV., vol. ED-32, No. 3, 3/85, By Koyanagi et al, pp. 562 et seq., "Optimum Design of n+-n-Double-Diffused Drain MOSFET To Reduce Hot-Carrier Emission".

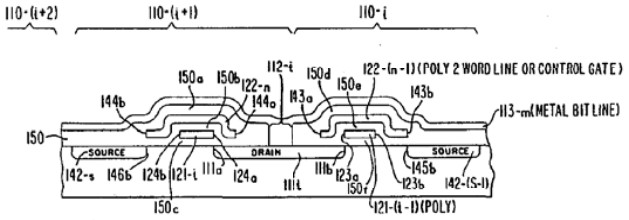
Primary Examiner—Martin H. Edlow

Attorney, Agent, or Firm—Skjerven, Morrill, MacPherson, Franklin & Friel

[57] **ABSTRACT**

A self-aligned split gate single transistor memory cell structure is formed by a process which self aligns the drain region to one edge of a floating gate. The portion of the channel underneath the floating gate is accurately defined by using one edge of the floating gate to align the drain region. The control gate formed over the floating gate controls the portion of the channel region between the floating gate and the source to provide split gate operation. The source region is formed sufficiently far from the floating gate so that the channel length between the source region and the closest edge of the floating gate is controlled by the control gate but does not have to be accurately defined.

8 Claims, 7 Drawing Sheets



KRNS000901

FIG. 1
PRIOR ART

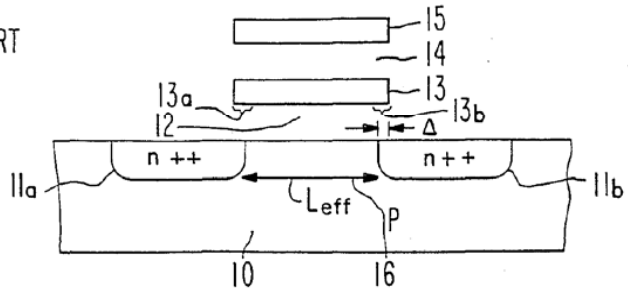


FIG. 2
PRIOR ART

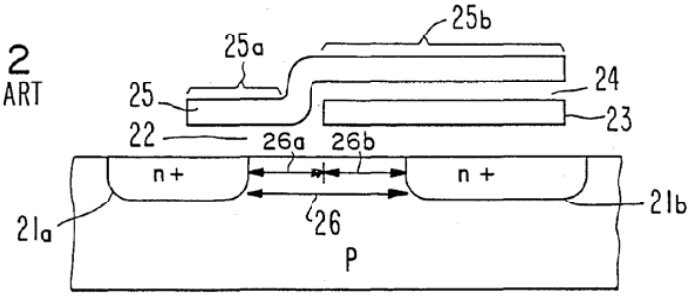


FIG. 3

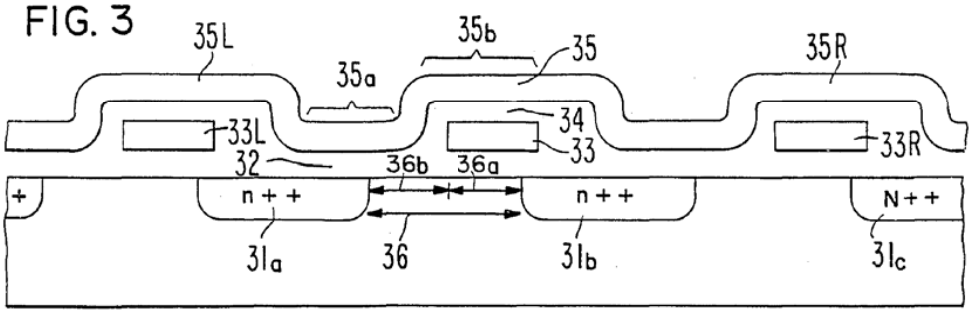
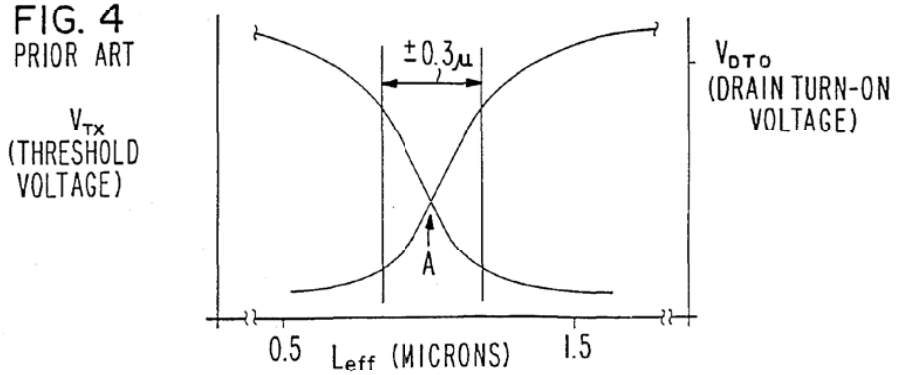


FIG. 4
PRIOR ART



KRNS000902

FIG. 5a

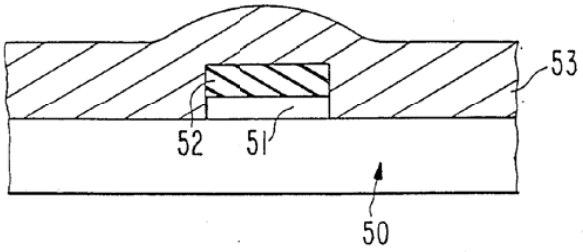


FIG. 5b

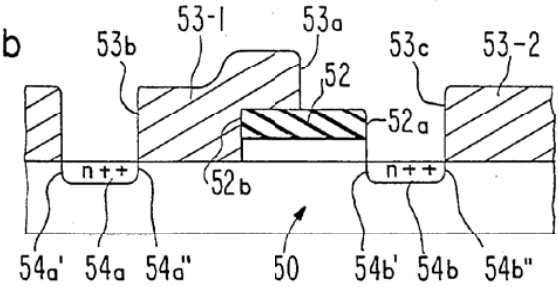


FIG. 6a

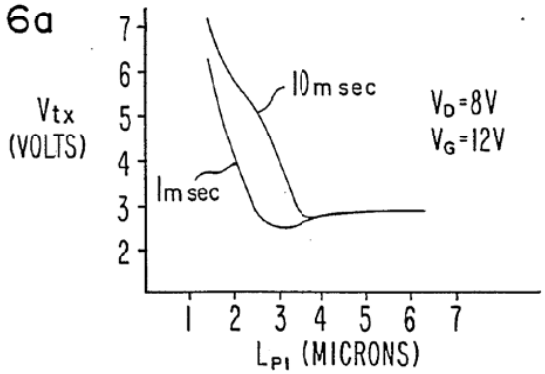
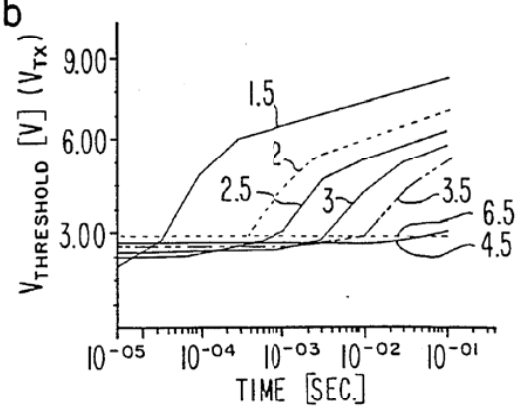


FIG. 6b



KRNS000903

FIG. 6c

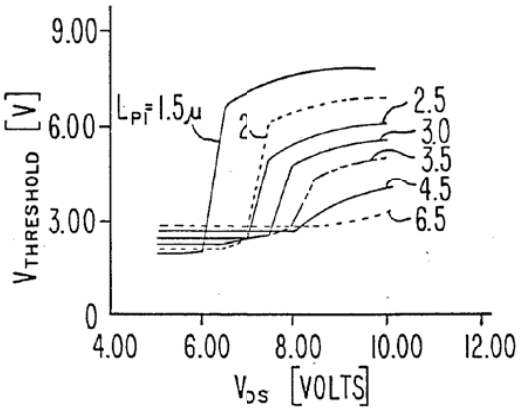
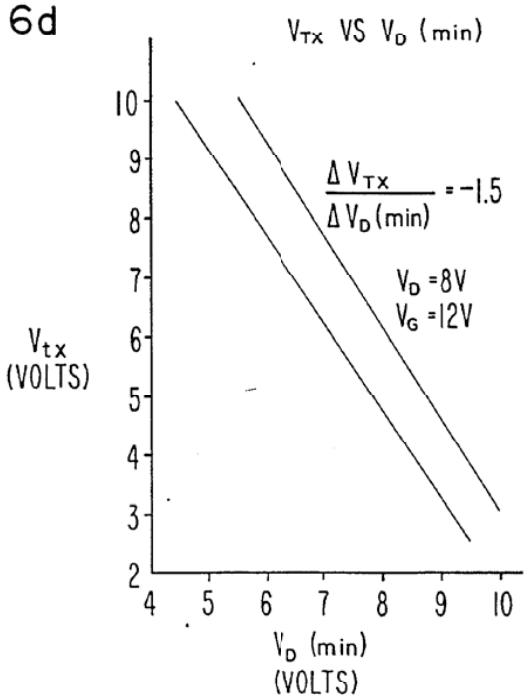


FIG. 6d



KRNS000904

FIG. 7a

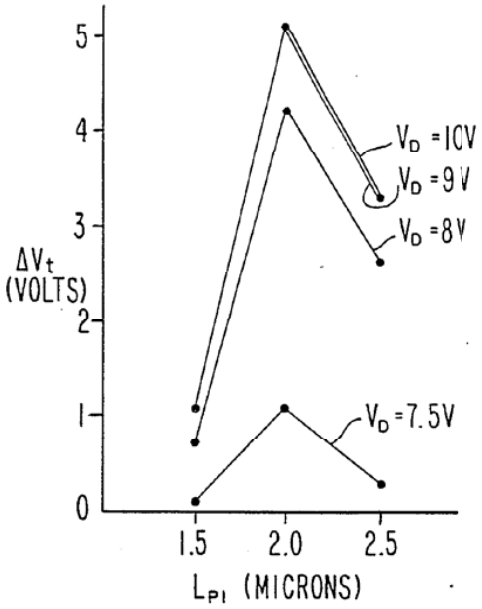


FIG. 7b

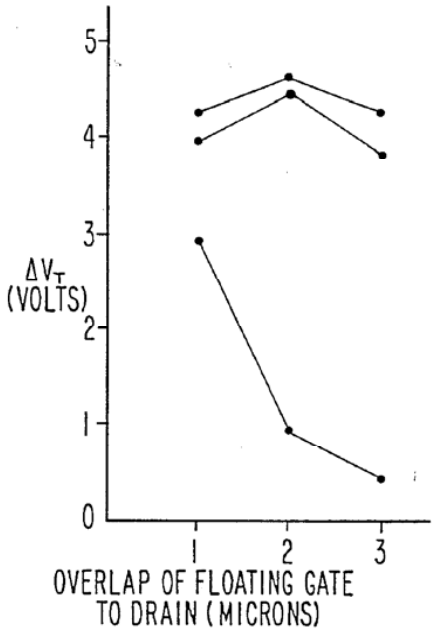
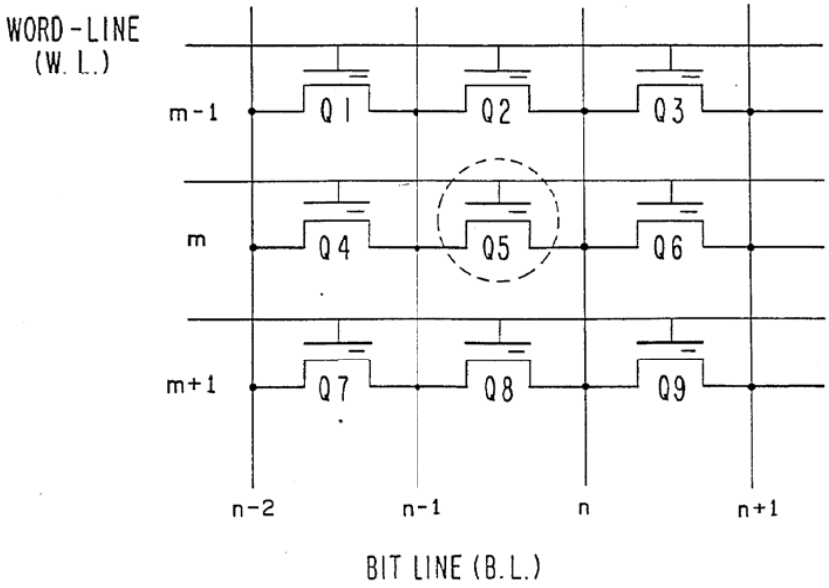


FIG. 8



KRNS000905

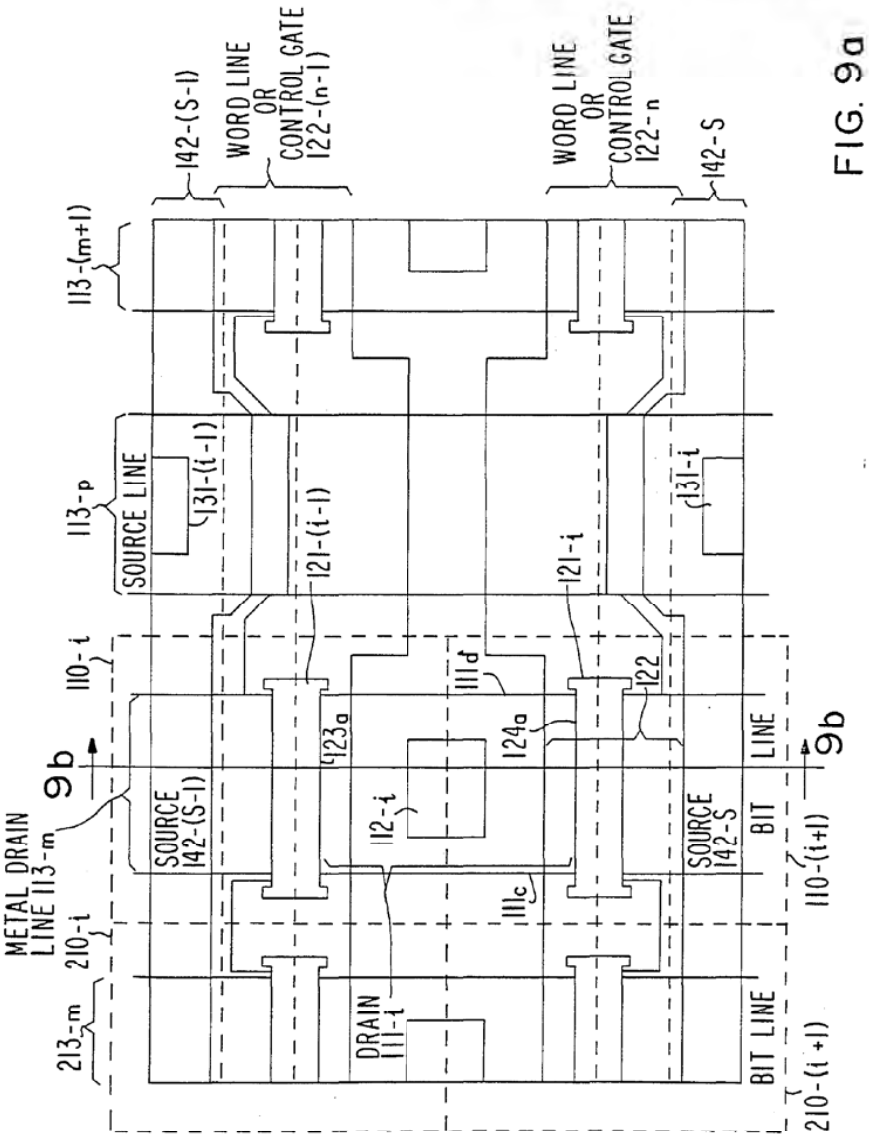


FIG. 9a

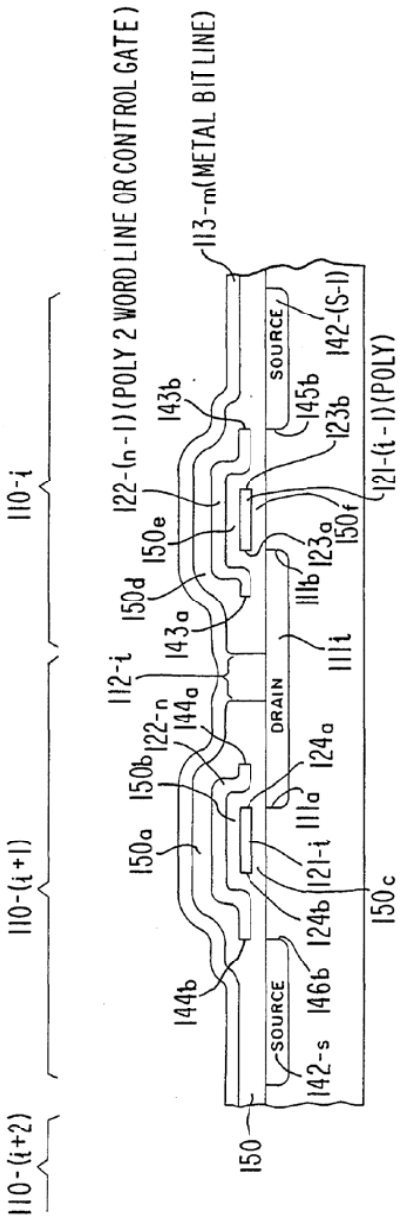


FIG. 9b

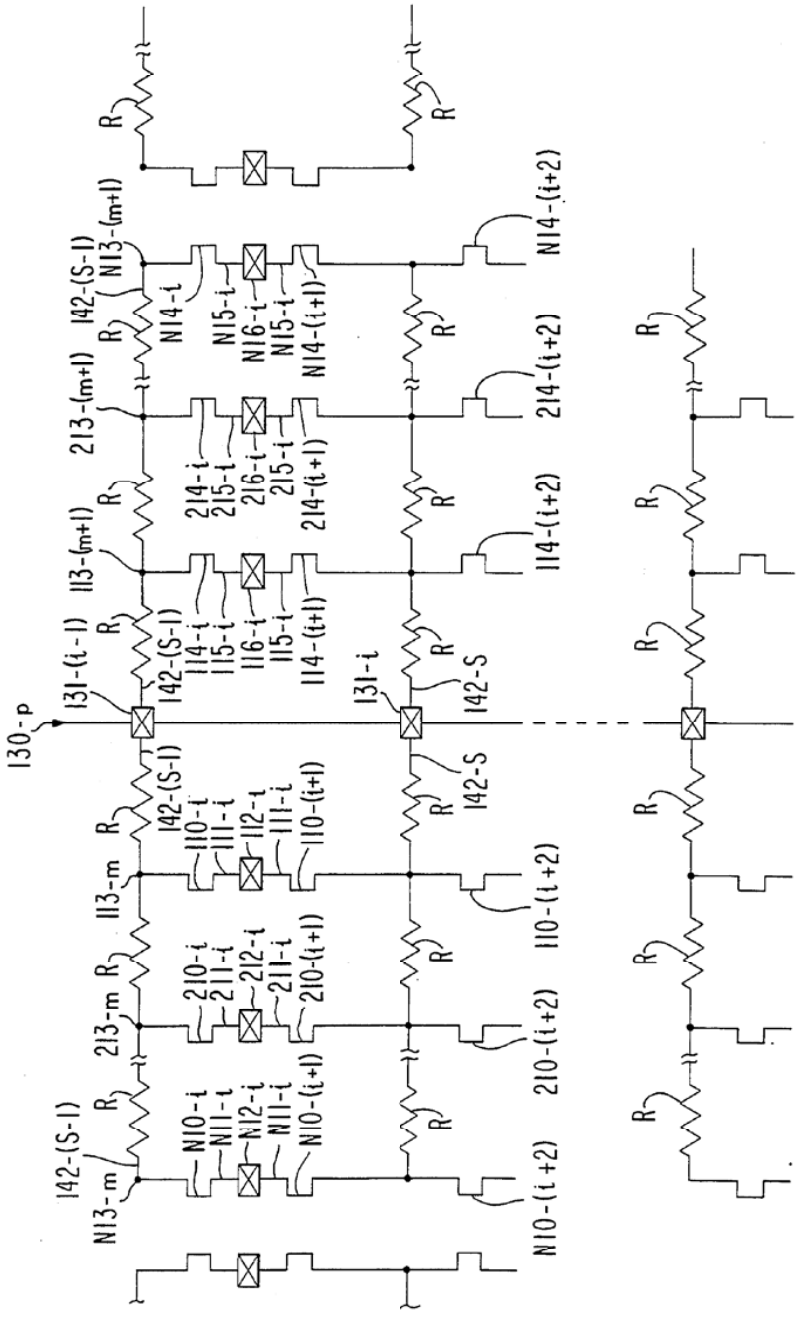


FIG. 9c

SELF-ALIGNED SPLIT GATE EPROM

RELATED APPLICATION

This application is a continuation-in-part of application Ser. No. 06/610,369, filed May 15, 1984, entitled "A SELF-ALIGNED SPLIT GATE EPROM", which application is assigned to Wafer Scale Integration, Inc. the assignee of this case.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a nonvolatile EPROM and more particularly to such an EPROM having a split gate (i.e., both a floating gate and a control gate) for controlling the writing and reading of each cell wherein the floating gate is self-aligned with the drain and the channel underlying the floating gate and the control gate is not self-aligned.

2. Prior Art

A split gate nonvolatile EPROM with increased efficiency 1B is disclosed in U.S. Pat No. 4,328,565 issued May 4, 1982 on an application of Harari, filed Apr. 7, 1980. As disclosed by Harari, the floating gate in an n channel EPROM cell extends over the drain diffusion and over a portion of the channel thereby to form a "drain" capacitance between 23 the drain and the floating gate and a "channel" capacitance between the channel and the floating gate. A control gate then overlaps the floating gate and extends over the remainder of the channel near the source diffusion thereby to form a "control" capacitance between the floating gate and the control gate. These three capacitances form the coupling for driving each cell. The inversion region in the channel directly under the control gate is established directly by a "write or read access" voltage applied to the control gate. The inversion region in the channel directly under the floating gate is established indirectly through the drain and control capacitances and the channel capacitance by the control gate voltage and by another write access voltage applied to the drain. A cell is erased either by ultraviolet illumination or by electrons from the floating gate tunneling through a region of thinned oxide. The nonsymmetrical arrangement of the control gate and floating gate with respect to source and drain allows a very dense array implementation. Other split gate structures are disclosed in an article by Barnes, et al. entitled "Operation and Characterization of N-Channel EPROM Cells", published in Solid State Electronics, Vol. 21, pages 521-529 B (1978) and an article by Guterman, et al. entitled "An Electrically Alterable Nonvolatile Memory Cell Using a Floating-Gate Structure", published in the IEEE Journal of Solid-State Circuits, Vol. SC-14, No. 2, April 1979.

FIG. 1 illustrates a typical EPROM of the prior art. In FIG. 1 a memory cell comprises n source region **11***a* and n++ drain region **11***b* separated by channel region **16**. Channel region **16** has an effective length L_{eff} as shown. Overlying channel region **16** is gate dielectric **12** on which is formed a floating gate **13**. Typically floating gate **13** is formed of polycrystalline silicon. Overlying floating gate **13** is insulation **14**, typically thermally grown silicon dioxide. Control gate **15** is formed above floating gate **13** on insulation **14**. The state of the transistor in FIG. **1** is determined by the charge placed on floating gate **13** When electrons are placed on floating gate **13**, the threshold voltage V_{tx} required on gate **15** to turn on the transistor (i.e., to form an n channel between source **11***a* and drain **11***b* thereby allowing current to flow from one to the other) is much greater than when no electrons are placed on floating gate **13**. As shown in FIG. **1**, regions **13***a* and **13***b* on floating gate **13** overlie the source **11***a* and drain **11***b*, respectively, by a small amount "Δ". Consequently, a capacitance is formed between the source **11***a* and floating gate region **13***a* and between the drain **11***b* and floating gate region **13***b*. If the overlap by gate **13** of the source **11***a* drain llb is the amount "Δ", then the capacitance C_{pp} between the floating gate **13** and the control gate **15** (both made of polycrystalline silicon) is given by the following equation:

$$C_{pp} \alpha A_{pp} \alpha W(L_{eff}+2\Delta_{FG,D}) \quad (1)$$

In equation 1, Cpp is the capacitance between the floating gate **13** and the overlying control gate **15** (this capacitance is proportional to A_{pp}) and A_{pp}, the area of the floating gate **13**, is just the width W of the floating gate B **13** (perpendicular to the sheet of the drawing) times the length of the floating gate **13** which is (L_{eff} $2\Delta_{FG,D}$)

The capacitance C_{PROM} between the floating gate **13** and the substrate **10** is proportional to the effective width W_{eff} (i.e. the width perpendicular to the sheet of the paper of the active area underneath the floating gate **13**) of the floating gate **13** times L_{eff}. Thus the capacitance C_{PROM} is

$$C_{PROM} \alpha A_{PROM} \alpha W_{eff}(L_{eff}) \quad (2)$$

The capacitive coupling $C_{FG,D}$ of the floating gate **13** to the drain **11***b* is given by

$$C_{FG,D} \alpha A_{FG,D} \alpha W_{eff}(\Delta_{FG,D}) \quad (3)$$

The coupling ratio $CR_{FG,D}$ of the capacitive coupling $C_{FG,D}$ of the floating gate **13** to drain **11***b* to the capacitive coupling C_{pp} of the floating gate **13** to the control gate **15** and the capacitive coupling C_{PROM} of the floating gate **13** to the substrate **10** is

$$CR_{FG,D} \alpha W_{eff}(\Delta_{FG,D})/[W_{eff}(L_{eff})+W\text{-}(L_{eff}+2\Delta_{FG,D})] \quad (4)$$

As L_{eff} becomes smaller and smaller the impact of the coupling of the drain on the performance of the PROM cell becomes greater and greater until in the limit, as L_{eff} becomes very, very small, this coupling approaches 0.3 (taking into account different oxide thicknesses and the difference between W and W_{eff}, for example). The overlay "Δ" depends on the process and is substantially fixed.

FIG. **2** shows the prior art split gate structure as illustrated by Harari in U.S. Pat. No. 4,328,565 issued May 4, 1982. The major concern in this structure relates to the length of portion **26***b* of channel **26** beneath floating gate **23** The structure of FIG. **2** is a nonself-aligned split gate structure. The total effective channel length **26** is defined by one mask and therefore is constant. Unfortunately, the length of the portion **26***b* of channel **26** beneath the floating gate **23** varies with mask alignment tolerances. Thus the effective channel length **26***b* depends strongly on the alignment process. As a result the best technology available today yields an effective

channel length tolerance **26***b* no better than ±0.5 to ±0.6 microns. For a typical nominal one micron effective channel length **26***b* the actual channel length will vary, due to manufacturing tolerances, over the range of about 1±0.6 micron. The result is a very wide variation in performance from one transistor memory cell to the next. Programming and read current are both very sensitive to channel length. Good cells will be perfect but bad cells will not work. A good device has an effective channel **26***b* (in one embodiment 0.8 microns) which lies between a too-short channel length (for example 0.2 microns or less, so that considering manufacturing variations there may not be any overlap at all of floating gate **23** over the channel **26** and thus there will be no programming of the cell) and a too-long channel length (for example greater than 1.4 microns) with unacceptably slow programming. The major issue in this prior art structure thus is the length of channel portion **26***b* (L_{eff}) rather than the coupling. Therefore in a structure such as that shown in FIG. **2** there can be coupling between drain **21***b* and floating gate **23** but if the channel length **26***b* is not carefully controlled, the memory cell is not going to perform as expected.

A major problem in the prior art EPROM of FIG. **1** relates to the relationship between the program threshold voltage V_{tx} and the drain turn on voltage V_{DTO} of the device. V_{DTO} is the voltage on the drain which, when capacitively coupled to the floating gate **13**, turns on the transistor. As shown in FIG. **4**, for L_{eff} as shown in FIG. **1** increasing from about 0.5 to 1.2 microns, the program threshold V_{tx} drops below the acceptable program threshold. On the other hand the drain turn-on voltage V_{DTO} becomes as high as the junction breakdown voltage for L_{eff} greater than about one micron. Below one micron, V_{DTO} is very low and may go as low as three to five volts which causes the array of EPROMS to fail. The crossover point is shown as "A" in FIG. **4**. In designing a regular EPROM, the crossover point A should be such that V_{tx} is high enough (i.e. greater than five volts) while VDTO is not too low (i.e. not lower than eight volts). However, both curves V_{DTO} and V_{tx} are quite steep at the crossover point A and thus the characteristics of the device are very sensitive to L_{eff}. So if the tolerance on Leff is even ±0.3 microns, which is very good, then the characteristics of the device are still relatively unpredictable. Obviously the desired solution is to eliminate the effect of V_{DTO} and optimize L_{eff} for V_{tx}.

SUMMARY OF THE INVENTION

In accordance with my invention, I overcome the problems of the prior art by providing a memory cell using a split gate structure containing both a control gate and a floating gate in which the floating gate is self-aligned to the drain region. The control gate is not self-aligned. "Self-aligned" means here that the portion of the transistor channel length under the floating gate will be defined by the floating gate itself regardless of any processing misalignments thereby insuring a constant channel length under the floating gate. To do this, a special process is employed wherein the floating gate is used to define one edge of the drain region. The source region is defined at the same time as the drain region but the alignment of the source region relative to the floating gate is not critical so long as the source region does not underlie and is spaced from the floating gate.

In a process in accordance with this invention, the diffused drain region (which also functions as a bit line and which corresponds to an elongated drain region of the type shown in the above-mentioned '565 patent) is formed using the floating gate to define one edge of the drain region. In the preferred embodiment, the drain and source regions are formed by ion implantation and one edge of the floating gate defines the lateral limit of one side of the drain region. A photoresist material partially extends over the floating gate in one direction and beyond the floating gate in the other direction and the source region is defined by an opening in the portion of this photoresist extending beyond the floating gate in the other direction. The result is to form a precisely defined channel portion L_{eff} of the channel region beneath the floating gate and a remaining relatively imprecisely defined portion of the channel region (to be controlled by a to-be-formed control gate electrode which is part of the word line) underneath the photoresist between the other edge of the floating gate and the source region.

In accordance with my invention, any misalignment between the.floating gate and the source region is covered by a to-be-formed control gate and has little effect on the operation of the memory cell while the floating gate is self-aligned to the drain region.

This invention will be understood in more detail in conjunction with he following drawings:

DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a prior art EPROM using a single floating gate beneath the control gate:

FIG. **2** illustrates the split gate structure of the prior art wherein the floating gate is not self-aligned to the drain region and the control gate is formed over part of the channel region:

FIG. **3** illustrates the split gate structure of this invention wherein the floating gate is self-aligned to the drain region and overlies but is insulated from an accurately defined portion L_{eff} of the channel region between the source and drain and the control gate overlies the floating gate and that portion of the channel region not overlain by the floating gate but is insulated therefrom;

FIG. **4** illustrates the relationship between threshold voltage V_{tx} and drain turn-on voltage V_{DTO} for the structure of FIG. **1**;

FIGS. **5***a* and **5***b* illustrate the novel process which I use to manufacture the novel self-aligned split gate structure of my invention;

FIGS. **6***a* through **6***d* illustrate the effect of the channel length L_{P1} under the floating gate on programming:

FIGS. **7***a* and **7***b* show the tight envelope of operation for the nonself-aligned structure and illustrate graphically the advantages of my self-aligned split gate structure: and

FIG. **8** shows in schematic form a memory array formed using the self-aligned split gate structure of my invention.

FIG. **9***a* shows the layout of a portion of a novel, high-speed EPROM incorporating the self-aligned split-gate structure of this invention;

FIG. **9***b* shows in cross-section two transistors from the structure of FIG. **9***a;* and

FIG. **9***c* shows schematically the architecture of one embodiment of the EPROM array portions of which are shown in FIGS. **9***a* and **9***b*.

DETAILED DESCRIPTION

The following detailed description is meant to be illustrative only and not limiting. Other embodiments of this invention will be obvious to those skilled in the art in view of the following description. In FIGS. 5*a* and 5*b*, only the cross-section of a single memory cell or a portion thereof is shown while in FIG. 3 two cells and part of a third are shown in cross-section. It should be understood that a semiconductor integrated circuit memory made in accordance with this invention employs a plurality of such cells together with peripheral circuits for writing data into memory and for accessing the data stored in the memory. For simplicity these circuits are not shown

The starting point for the process of my invention to yield my novel self-aligned split gate structure is the same as in the nonself-aligned split gate structure of the prior art and in particular of the '565 patent. Thus as shown in FIG. 5*a* a silicon substrate 50 typically having a resistivity of 10–50 ohm-centimeters has formed thereon in a standard manner a layer of gate oxide **51**. Gate oxide **51**, typically 300 angstroms thick, then has formed on it a first layer of polycrystalline silicon (often called "poly 1") which is patterned as shown in FIG. 5*a* to to form floating gate 52 a floating gate 52. The oxide **51** beneath the portions of polycrystalline silicon removed to form floating gate **52** is then removed by an etching process (typically a plasma etch) and a photoresist layer **53** is then formed over the top surface of the structure.

As shown in FIG. 5*b*, photoresistor layer **53** is then patterned so that a particular segment **53-1** of photoresist is formed to partially overlie floating gate **52**. Photoresist segment **53-1** has a right edge **53***a* which is formed to overlie the floating gate **52** somewhere near its middle and a left edge **53***b* which is formed to the left of left edge **52***b* of floating gate **52**. The length of floating gate **52** is typically 1.5 to 2 microns and thus it is not difficult to ensure with sufficient certainty given typical tolerances in the manufacturing process that edge **53***a* is to the left of right edge **52***a* of floating gate **52** even for a reasonably expected worst case mask misalignment during the manufacturing process. It is also quite simple to insure that left edge **53***b* is sufficiently to the left of left edge **52***b* of floating gate **52** so that left edge **52***b* of floating gate **52** is never exposed, even in a worst case alignment mismatch of masks during manufacture. Thus the to-be-formed source **54***a* will always be laterally spaced from the left edge **52***b* for floating gate **52**.

Following the formation of patterned photoresist segment **53-1** the structure is subjected to an ion implantation at a selected well-known dosage (typically 4×10^{12} per cm^2) to form n++ drain region **54***b* and n. source region **54***a* in the top surface of the semiconductor material **50**. The region **54***b* has its left edge **54***b'* defined by the right edge **52***a* of floating gate **52** and its right edge **54***b''* defined by the left edge **53***c* of patterned photoresist **53-2**. The source region **54***a* has its right edge **54***a''* defined by the left edge **53***b* of patterned photoresist segment **53-1** . Thus the drain region **54***b* is self-aligned to the right edge **52***a* of floating gate **52**. However, the right edge **54***a''* of source region **54***a* is self-aligned to the left edge **53***b* of photoresist segment **53-1**. The uncertainty in the location of left edge **53***b* of patterned photoresist segment **53-1** relative to left edge **52***b* of floating gate **52** represents an uncertainty in the length of the control gate channel (corresponding to channel portion **36***b* in FIG. 3) and not of the floating gate channel L_{eff} (corresponding to channel portion **36***a* in the center cell of FIG. 3). By placing a proper voltage on the to-be-formed control gate, (corresponding to control gate **35** in FIG. 3), the channel length under the control gate becomes irrelevant and the conduction or nonconduction of the total channel is determined by the voltage placed on the floating gate **52** (corresponding to floating gate **33** in FIG. 3). Because floating gate **52** is uniformly coupled to drain **54***b* in all transistors in a memory array made in accordance with this invention and further because the effective channel length L_{eff} (corresponding to channel **36***a* in FIG. 3) is substantially the same underneath all floating gates **52** in all transistors in a memory array formed in accordance with this invention, the structure of this invention yields a split gate programmable EPROM capable of being manufactured with much higher yield than the prior art EPROMs.

The remaining steps in the process are standard well known steps in the silicon gate EPROM technology. Insulation (not shown) is formed over floating gate **52**. A control gate (often called "poly 2") corresponding to control gate **35** in FIG. 3 is formed, usually as part of a word line. The resulting structure appears as shown in FIG. 3. FIG. 3 shows floating gate **33** and floating gates **33L** and **33R** formed to the left and right of floating gate **33**. All three floating gates have their right edges self-aligned to the left edges of the underlying drain regions in accordance with this invention. The drain region for a given cell doubles as the source region for the cell to the right.

The finished structure made by the process of this invention as illustrated in FIGS. 5*a* and 5*b* is shown in FIG. 3. In FIG. 3 the floating gate **33** has been formed prior to the formation of the source and drain regions **31***a* and **31***b*. The floating gate **33** is formed on a thin layer of insulation overlying a portion of the to-be-formed channel region between the source and drain. The right edge of the floating gate **33** has been used to define one edge of the drain region **31***b*. Overlying floating gate **33** is insulation **34** (typically silicon oxide) and overlying oxide **34** is the control gate **35**. A portion **35***a* of control gate **35** overlies a second portion of the channel region between the left end of the floating gate **33** and the source region **31***a*. As described herein, the channel region **36***b* beneath portion **35***a* of control gate **35** can have a length **36***b* which varies substantially without affecting the performance of the device.

The structure shown in the central part of the cross-section in FIG. 3 is but one cell of a plurality of such cells. In a typical virtual ground structure the drain **31***b* for the cell shown in cross-section in FIG. 3 serves as the source for another cell located just to the right. Likewise the source **31***a* serves as the drain for a second cell located just to the left. The portions of the floating gates **33L** and **33R** associated with these adjacent cells are shown in FIG. 3.

Note that the floating gate **52** (FIGS. 5*a* and 5*b*) becomes capacitively coupled to drain region **54***b* by the lateral diffusion of left edge **54***b'* beneath floating gate **52** during further processing of the structure. This lateral diffusion is typically around 0.3 microns. However contrary to the prior art, the floating gate **52** is formed before the formation of the drain region **54***b*, rather than after, and is precisely self-aligned to one edge of the drain region **54***b*.

FIG. 6*a* illustrates the variation in threshold voltage versus the drawn channel length (L_{P1}) of the floating gate ("poly 1"). In FIG. 6*a* the ordinate is the program threshold and the abscissa is the length of the floating gate channel L_{P1} in microns. (Of importance, FIGS. 6*a* through 6*d* and 7*a* and 7*b* use drawn dimensions. However, the channel lengths 36*a* and 36*b* shown in FIG. 3 are the effective dimensions after processing. Thus channel length 36*a* is denoted Leff to represent the effective length of this channel after processing, while before processing this channel length is a drawn dimension and as such is denoted by the symbol L_{P1}. Accordingly, each of the dimensions L_{P1} shown in FIGS. 6*a* through 6*d* and 7*a* and 7*b* must be corrected (i.e., reduced) by a given amount (approximately 0.5 microns), to reflect the effect of processing. Naturally the amount of the correction will vary with the processing.) The threshold voltage V_{tx} obtained or programmed in a given time for a given drain voltage and gate voltage (corresponding in FIG. 6*a* to a drain voltage of 8 volts and a gate voltage of 12 volts) drops rapidly as the length of the channel LP1 under the floating gate 52 (FIG. 5*b*) increases to a minimum V_{tx} of about 2.5 volts for L_{P1} of somewhere between 3 to 4 microns and then increases slightly. This minimum V_{tx} corresponds to the initial device threshold before programming. The threshold V_{tx} represents the voltage which must be applied to the control gate (such as gate 35 in FIG. 3) to turn on the transistor beneath the control gate as shown in FIG. 3 when the cell containing that transistor has been programmed Thus as the length of the channel 36*a* underneath the floating gate 33 increases (FIG. 3) the threshold voltage necessary to turn on the transistor and create a channel from the source region 31*b* to the drain 31*a* decreases. As is shown in FIG. 6*a*, both 1 millisecond and 10 millisecond programming times yield substantially the same shaped curve.

FIG. 6*b* illustrates the effect of the length of the channel 36*a* underneath floating gate 33 on the threshold voltage (ordinate) versus programming time (abscissa). The various curves reflect different lengths L_{P1} of the channel 36*a* (FIG. 3) beneath floating gate 33 in microns. As these channel lengths increase, the threshold voltage for a given programming time drops. Thus for a programming time of 10^{-2} seconds, the threshold voltage for a 1.5 micron channel length L_{P1} is approximately 7 volts whereas the threshold voltage for a 3.0 micron channel L_{P1} is about 4 volts. These curves were obtained for a voltage VDS from the drain to the source of 8 volts and a voltage on the control gate 35 of 12 volts. The curves of FIG. 6*b* illustrate that the shorter the floating gate the stronger the field which is formed and therefore the greater the number of electrons which are placed on the floating gate thereby resulting in a larger threshold voltage V_{tx} to turn on the transistor.

FIG. 6*c* is a plot of threshold voltage V_{tx} (ordinate) versus the voltage on the drain 31*b* (FIG. 3) with the length of channel 36*a* beneath floating gate 33 as the parameter on the various curves. For a given drain voltage V_{DS} (for example 8 volts) the threshold voltage V_{tx} goes up as the length L_{P1} of the channel 36*a* beneath floating gate 33 goes down. The curves of FIG. 6*c* were taken with a control channel L_{P2} (corresponding to the drawn dimension of channel 36*b* in FIG. 3) beneath the control gate 35 of 2.5 microns, a gate voltage on control gate 35 of 12 volts and a programming time of 10 milliseconds (10^{-2} seconds). These curves illustrate that once a given drain voltage difference V_{DS} is achieved between the drain and the source, increasing the drain voltage beyond a given amount has substantially little effect on the threshold voltage V_{tx} of the transistor. In other words, $\Delta V_{tx}/\Delta V_{DS}$ becomes substantially zero thereby showing that increasing the drain voltage coupled to the floating gate has little effect on the programming of the transistor. Thus after the program threshold voltage V_{tx} is reached increasing the drain to source voltage V_{DS} does not achieve any significant improvement in performance.

As L_{P1} increases, the threshold voltage V_{tx} at which ΔV_{tx} over ΔV_{DS} becomes very small decreases. So increasing V_{DS} does even less for structures with longer floating gates.

In FIG. 6*c* each consecutive point on a given line for a given L_{P1} represents an additional 10 milliseconds of programming time rather than just 10 milliseconds of programming time. Accordingly the curves for V_{tx} versus V_{DS} in FIG. 6*c* would be even flatter than shown in FIG. 6*c* if a constant programming time was applied to program the cell from different V_{DS} start points.

FIG. 6*d* illustrates the very tight predictability of threshold voltage V_{tx} versus V_D (min) for the structure of this invention. V_D (min) is defined as the minimum VDS needed to start programming (i.e., to start efficient electron flow onto the floating gate). In FIG. 6*c* V_D(min) is the V_{DS} at which the curve shows a break point sharply to the right. This break point or "knee" corresponds to the V_D(min) plotted in FIG. 6*d*.

The relationship of FIG. 6*d* to FIG. 6*c* illustrates a basic print of my invention. In a 256 K EPROM the time to program the cells in the EPROM theoretically equals 256 K times the time to program each cell divided by 8 (ROMs are programmed one byte at a time). Therefore, if the programming time of each cell can be significantly reduced, the efficiency of programming a large number of EPROMs can be proportionally increased. I have discovered that to program to a given threshold voltage V_{tx} in a given programming time, the key is to control the length of L_{P1} and in particular to make this length (which is related to the channel 36*a* in FIG. 3) as small as practical without generating punch through from the source to the drain. As shown by analysis of FIG. 6*d*, the threshold voltage V_{tx} is increased for a given programming time by decreasing V_D(min). As shown in FIG. 6*c* V_D(min) decreases as L_{P1} decreases in length. Accordingly, decreasing L_{P1} is the key. to programming to a given threshold voltage V_{tx} in a given time. My invention not only allows a small effective channel length L_{eff} to be achieved beneath the floating gate but allows this channel length to be achieved in a controllable and reproducible manner throughout an EPROM array thereby to obtain repeatable and consistent results throughout the array.

FIG. 7*a* illustrates change of threshold voltage, ΔV_T for three different valves of L_{P1} (i.e., three different drawn channel lengths beneath the floating gate) for the structure shown in FIG. 2. In a nonself-aligned structure, the proper length of the channel under the floating gate is crucial to achieve maximum threshold voltage V_{tx}. As shown in FIG. 7*a* if the channel length 36*a* becomes too short (for example, 1.5 microns), then punch-through occurs between the source 31*a* and drain 31*b* during programming resulting in a failure to program the device. The proper alignment of a floating gate in the nonself-aligned structure to optimize the length of the channel 36*a* beneath the floating gate 33 and the overlap of the floating gate to drain is crucial.

The very sharp peak in FIG. 7*a* reflects the variation in V_{tx} with channel length L_{Pl}. FIG. 7*a* shows that to optimize the device for the minimum channel length L_{Pl} in terms of programming efficiency results in a lower initial threshold before programming and higher final threshold after programming so as to obtain a higher read current. This means a lower impedance in the circuit which in turn means that during read a capacitor in the sense amplifier in the peripheral circuitry of the memory discharges faster through a programmed transistor than otherwise would be the case resulting in shorter access time. Three effective channels beneath the floating gate (1.5 micron, 2.0 micron and 2.5 micron) are shown in FIG. 7*a*. The parameter ΔV_T (representing the change in threshold voltage as a function of different channel length) is illustrated by the curves. This change in voltage is particularly pronounced as one goes from 1.5 to 2 to 2.5 micron length for L_{Pl}. The change in V_{tx} as a function of channel length is similar to that shown in FIG. 6*a* for the self-aligned structure of my invention. However, as one goes from a 2 micron L_{Pl} to 1.5 micron L_{Pl} and shorter, a new phenomenon appears reflecting possible punch through from the source to the drain and Vtx thus is lower than would be expected. The nonself-aligned curve shows that a proper L_{Pl} is critical to obtaining a predicted threshold voltage. However, with nonself-aligned floating gate technology L_{Pl} can vary even across a given chip causing a variation in V_{tx} from cell to cell within a given memory. Often this variation is unacceptable. As can be seen by the curves of FIG. 7*a*, a given memory can have L_{Pl} from cell to cell varying for example from 1.5 microns all the way to 2.5 microns.or greater because of misalignment in the masking during the processing of the wafer. Accordingly, V_{tx} is unpredictably variable across the wafer often resulting in unacceptable performance.

FIG. 7*b* shows the effect of overlap and V_D on threshold voltage. For the nonself-aligned device the structure must be aligned so that the 3 sigma worst case of alignment gives a satisfactory channel length **36***a* beneath floating gate **33**. Increasing the coupling between the floating gate and the drain does not improve the threshold voltage of the device for given programming conditions so overlapping the drain with the floating gate does not help. The more overlap of the floating gate to the drain means the more electrons required to charge the floating gate for a given channel length **36***a* beneath the floating gate. So instead of improving the efficiency of the device, increasing the overlap of the floating gate to the drain actually decreases this efficiency. A minimum overlap of the floating gate to the drain is needed to insure that accelerated electrons hit and lodge in the floating gate rather than in the control gate or the word line.

FIG. 7*b* shows that as the overlap of the nonself-aligned structure increases, the ΔV_T actually declines for a given V_D. Again, this shows that the coupling between the drain and the floating gate is not helpful to achieving a desired V_{tx} and indeed can even be harmful.

The circuit of this invention is highly scaleable and retains its self-aligned character as it is scaled.

An important effect of this invention is that by choosing the correct L_{Pl} the programming time for a memory array can be substantially reduced. For example, a prior art 256 K EPROM takes approximately 150 seconds or 2½ minutes to program. A 256 K EPROM using the structure of this invention can be programmed in approximately 30 seconds. This is a substantial improvement resulting in lower programming costs and lower test costs.

An additional advantage flowing from this invention is that the uncertainty in the location of the floating gate due to mask alignment tolerances is substantially reduced compared to the uncertainty in the location of the floating gate in the prior art nonself-aligned structure and in the standard prior art EPROM (nonsplit gate but self-aligned). Table 1 illustrates this improvement with respect to the self-aligned split gate structure of this invention compared to the standard non-split gate self-aligned structure of the prior art.

TABLE I

	Standard EPROM (Non-split but self-aligned gate)	Self-aligned split gate structure of this invention
STEP 1	Poly 1 (Floating gate) Critical dimension not defined but non-critical dimensions are defined	Poly 1 (Floating gate) Critical dimension defined
STEP 2	Poly 2 (Control gate) Define critical dimensions of control gate Structure - Accuracy degraded because of rough, nonplanar topology associated with two layers of polycrystalline silicon	-
STEP 3	Poly 1 critical dimension defined using Poly 2 as a mask	

Table I compares only the critical steps in the two processes used to define the floating gate and thus the crucial channel length L_{eff}. L_{eff} is the important channel length in the self-aligned split gate structure of this invention and in any EPROM structure. Note that in a standard non-split gate self-aligned structure L_{eff} is the total channel length between the source and drain.

As shown in Table I three steps are required to define the critical dimension of the floating gate in the standard non-split gate self-aligned structure. In the first step only the noncritical dimensions corresponding to the width (but not the length) of the channel beneath the floating gate are defined. The critical dimensions of the floating gate corresponding to the channel length beneath the floating gate are not defined. In step 2 the second layer polycrystalline silicon from which the control gate will be fabricated is deposited. The critical dimension of this second layer (known as "poly 2") is defined in step 2. This dimension corresponds to the channel length between the to-be-formed source and drain regions. However, the accuracy with which the critical dimension of the control gate is fabricated is degraded because of the rough nonplanar topology associated with the two layers of polycrystalline silicon deposited on the wafer. In the third step the first layer of polycrystalline silicon (poly 1) has its critical dimension (corresponding to channel length L_{Pl}) defined using the second layer of polycrystalline silicon as a mask. Again, the accuracy with which the critical dimension of the first layer of polycrystalline silicon is defined is degraded due to the uneven topology of the structure.

In contrast, the self-aligned split gate structure of my invention defines the critical dimension of the poly 1 floating gate layer in step 1.

As the above comparison shows, the channel length L_{PI} for the standard nonsplit gate self-aligned structure is equal to the drawn length of the channel plus or minus the uncertainty in the critical dimension associated with the poly 2 definition step plus or minus the uncertainty introduced in the critical dimension of the channel length associated with poly 1 using poly 2 as a mask. Thus the uncertainty in the effective channel length in the standard nonsplit gate self-aligned structure has two components introduced by two critical dimensions. On the other hand, using the self-aligned split gate structure of my invention, only one uncertainty in a critical dimension occurs and that occurs in the first step where the poly 1 critical dimension is defined and the topology is smooth. Accordingly my invention yields a double processing advantage over the process by which the standard non-split gate self-aligned structure of the prior art is made by eliminating one critical dimension in defining L_{eff} and by introducing a much smoother topology during the formation of the critical channel length L_{eff}.

Table 2 compares the critical steps required to define the poly 1 floating gate in the nonself-aligned split gate structure of the prior art compared to the single step required to define the floating gate in the self-aligned split gate structure of my invention.

TABLE II

	Nonself-aligned split gate structure	Self-aligned split gate structure of this invention
STEP 1	Source and Drain Implanted	Poly 1 (Floating Gate) Define critical dimension
STEP 2	Poly 1 (Floating Gate) Define critical dimension	

Step 1 in fabricating the prior art nonself-aligned split gate structure is to implant the source and drain regions in the device. Step 2 is then to deposit the poly 1 layer and then form the floating gate from this layer. The critical dimension L_{PI} is defined by this step. Unfortunately, uncertainty in the length of L_{PI} results from the uncertainty in the critical dimension of the poly 1 plus or minus the misalignment of the mask used to define the critical dimension of the floating gate relative to the underlying drain region. Typically the uncertainty in the critical dimension is +0.3 microns while the uncertainty due to the mask misalignment is +0.6 microns. When combined in a statistical sense (root means square) the total uncertainty in L can be +0.6 or +0.7 microns. To the contrary, using the self-aligned split gate structure of my invention, the critical dimension of the poly 1 floating gate is defined with an uncertainty at most of about +0.3 microns. Accordingly, my invention achieves a substantial improvement in manufacturing accuracy over the prior art nonself-aligned split gate structure.

FIG. **8** illustrates an EPROM array fabricated using the self-aligned split gate structure of my invention. For simplicity, an array of nine (9) transistors or cells is shown. The programming and reading of cell or transistor Q5 will be described. Note that the array comprises word line rows m−1, m and m+1 and bit line columns n−2, n−1, n and n+1. Column n−2 is the source of transistors Q1, Q4 and Q7 while column n−1 is the drain of transistors Q1, Q4, and Q7 and the source of transistors Q2, Q5 and Q8. Similarly, column n is the drain of transistors Q2, Q5 and Q8 and the source of transistors Q3, Q6 and Q9. Column n+1 is the drain of transistors Q3, Q6 and Q9.

In operation, to read device m,n (i.e. cell Q5) all bit lines except n−1 are set at 2 volts. Bit line n−1 is set at ground. Word line m is set at 5 volts while all other word lines except m are set at ground.

To program device m,n (i.e., cell Q5) all bit lines except n are set at ground while bit line n is set at 8 or 9 volts. All word lines except m are set at ground while word line m is set at 12 volts. During programming, device m, n+1 (i.e., cell Q6) is also in programming condition but in the reverse configuration (i.e., the high voltage is applied away from the floating gate). In this configuration there is no programming of m, n+1. This asymmetry in the split gate EPROM is what enables one to utilize the virtual ground approach.

An additional embodiment of this invention is illustrated in FIGS. *9a*, *9b* and *9c*. FIG. *9a* illustrates in top view the layout of an embodiment of this invention which decreases the switching time necessary to read the state of a cell. Naturally, to increase the speed of a memory the time necessary to read the state of each cell in the memory should be decreased. The smaller the cell current or the larger the capacitance associated with the bit line connected to a given cell, the longer it takes to read the state of the cell. In the previously described embodiments of this invention, the drain region for one cell may serve as the source region for another cell. Thus when a cell is being read the source may serve as virtual ground. To read one cell in a memory and then read a second cell, the drain region of the second cell which has previously served as virtual ground must be switched to a higher voltage. The time necessary to do this depends upon the capacitance of the drain region. To do this more rapidly, the drain capacitance (also called the bit line capacitance) must be reduced. The structure of FIG. *9a* does this by using a novel array architecture utilizing the self-aligned split gate EPROM of this invention. Instead of a virtual ground which can function as both a source and a drain region and which achieves the high cell density as in the embodiment described above in conjunction with FIGS. **3**, *5a* and *5b*, the structure of FIG. *9a* uses a solid, dedicated source line **130**-p as in a standard EPROM and dedicated bit lines (such as bit lines **213**-m, **113**-m and **113**-(m+1). The source line is not switched from virtual ground to a high voltage but rather is always kept at a voltage near the level at which the cell is to be switched. By doing this the switching time can be decreased. The source line **130**-p comprises a metal line formed on insulation over the array in a direction orthogonal to source regions **142**-(s-1) and **142**-s. Vias **131**-(i-1) and **131**-i connect metal source line **130**-p to source regions **142**-(s-1) and **142**-s, respectively. Source regions **142**-(s-1) and **142**-s are typically formed by ion implantation. Formed orthogonal to source line **130**-p are word lines **122**-n and **122**-(n-1). Word lines **122**-n and **122**-(n-1) function as control gates and are formed over but separated on insulation from floating gates **121**-i and **121**S3 (i-1). Floating gates (of which gates **121**-i and **121**-(i-1) are shown in top view in their entirety) are formed of polycrystalline silicon overlying but insulated from the channel region between an underlying drain and source. In accordance with the description given above in conjunction with the structure of FIGS. **3**, *5a* and *5b*, drain **111**-i is formed by ion-implantation using edges **123**-a and **124**-a of floating

gates 121-(i-1) and 121-i, respectively to define the top and bottom edges 111*b* and 111*a*, respectively (FIG. 9*b*) of drain 111-i. Sides 111*c* and 111*d* of drain 111-i are bounded by oxide isolation. Thus each drain region such as region 111-i is disconnected from the other similar drain regions.

As will be shown shortly, the width of the source regions 142-(s-1) and 142-s can be reduced substantially by forming each source region (such as source regions 142-s and 142-(s-1)) by ion implantation using one edge of each word line (such as word line 122-n or 122-(n-1) respectively) as a mask to define the edge of the corresponding source region during the n+ ion implantation used to form the n+ regions of the peripheral access and logic transistors on the EPROM.

The main difference in the structure shown in FIG. 9*a* using the self-aligned split gate embodiments shown in FIGS. 3, 5*a* and 5*b*, and prior art EPROM arrays is the location of the bit lines 213-m, 113-m and 113-(m+1) and the word lines 122-(n-1) and 122-n in terms of layout. The advantage of the structure of FIG. 9*a* is that the bit line 113 associated with a given cell 110 does not have to be switched all the way from ground to a voltage necessary to detect the state of this cell but is always held at a voltage close to the read voltage. The selection of the particular cell to be read is done by the word line 122. A second advantage is that a bit line 113-m, 213-m . . . has much less capacitance than a typical prior art bit line. The reason for this is that while the bit line (such as bit line 113-m) is connected to a plurality of drain regions (such as drain 111-i of cell 110-i+1) arranged in a column, each drain such as drain 111-i functions as the drain in only two adjacent transistors (transistors 110-(i+1) and 110-i as shown in FIG. 9*a*) and each drain is not connected as part of a continuous diffused (or in this case, ion-implanted) region to the other drain regions. Thus the capacitance associated with each drain region 111 is reduced compared to the capacitance associated with a continuous drain diffusion of the prior art.

As shown in FIG. 9*a* and in cross-section in FIG. 9*b*, the ion-implanted drain region 111-i serving memory cell 110-i+1 is contacted by a via 112-i formed in insulation 150 over the drain region 111-i. Metal bit line 113-m (also called a metal drain line) electrically contacts drain 111-i through via 112-i. Formed directly adjacent to drain 111-i and self-aligned with drain 111-i as described above in conjunction with FIGS. 3, 5*a* and 5*b*, are two floating gates 121-(i-1) and 121-i formed of a first layer of polycrystalline silicon ("poly 1"). Overlying floating gate 121-i is a control gate 122-n (also called a word line) formed of a second layer of polycrystalline silicon ("poly 2"). The second layer of polycrystalline silicon 122-n extends the length of 2N transistors in the array to form a word line (also called a control gate) and is orthogonal, in the embodiment shown, to both metal drain line 113-m and metal source contact line 130-p. N is an integer which in accordance with this invention is preferrably 2, 4 or 8.

The cell 110-(i+1) includes part of a source diffusion denoted as 142-s in FIG. 9*a*. The source diffusion 142-s for cell 110-(i+1) is also the source diffusion for cell 110-(i+2) (not shown in FIG. 9*a* but shown schematically in FIG. 9*b* and 9*c*) just as the drain diffusion 111-i is the drain diffusion-for cell 110-i as well as for cell 110-(i+1). Source diffusion 142-s also serves as the source region for other cells in a given row. Thus each source diffusion except the first actually serves as the source region for 4N memory cells. The polycrystalline silicon word line 122-n is formed with substantial overlap over the source and drain to prevent misalignment from affecting the ability of a given cell to turn on when read. Edges 123*a* and 124*a* of the poly-1 floating gates 121-(i-1) and 121-i serve, as described above in conjunction with FIGS. 3, 5*a* and 5*b*, to define the edges 111*b* and 111*a*, respectively, of drain region 111-i as illustrated in FIGS. 9*a* and 9*b*.

FIG. 9*b* also illustrates the symmetrical structure memory cells 110-i+1 and 110-i of this invention. These two cells share a common drain region 111-i as shown. Control gate 122-n (typically formed of polycrystalline silicon) has a left edge 144*b* which extends over source region 142-s. Source 142-s serves as a source not only for cell 110-(i+1) but also for cell 110-(i+2) (adjacent and below cell 110(i+1) in FIGS. 9*a* but not shown in FIG. 9*a*). Overlying source region 142-s is a metal contact layer 130-p (FIG. 9*a*) which is connected to the source line 142-s by a contact 131-i.

In fabricating the structure of FIGS. 9*a*, 9*b* and 9*c* it is clear that the source region 142-s, for example, can be fabricated using left edge 144*b* of polycrystalline word line 122-n as a guide to define right edge 146*b* of source region 142-s. When this is done, the width of control line 122-n and all similar control lines can be reduced by the tolerance otherwise placed on this width to insure that it is properly aligned over source region 142-s. Thus a saving in space can be achieved using this technique of at least half a micron in width of control line 122-n. When similar savings are made in conjunction with control line 122-(n-1) by self-aligning left edge 145*b* of source 142-(s-1) with the right edge 143*b* of control line 122-(n-1), a substantially smaller array can be achieved. Advantageously, sources 142 are formed using the n+ ion implant used to form the MOS transistors in the logic and access circuitry in the peripheral regions of the memory array.

FIG. 9*c* illustrates schematically the layout of an array utilizing the structure shown in FIGS. 9*a* and 9*b*. As shown in FIG. 9*c*, metal source line 130-p is shown extending vertically down the center of the array. Contacts 131-i and 131-(i-1) are shown schematically to illustrate the vias through the underlying insulation through which metal line 130-p electrically contacts the laterally extending source regions 142-s and 142-(s-1). As described above, source regions 142-s and 142-(s-1) are preferrably formed by ion implantation. Each lateral source region serves as the source for up to 2N transistors on each side of metal source line 130-p where N typically is an integer and can be 2, 4 or 8. Metal bit lines 113-m, 213-m . . . to N13-m are shown to the left of metal source line 130-p and extending parallel to metal source line 130-p while metal bit lines 113-m+1, 213-(m+1), . . . to N13-(m+1) are shown to the right of metal source line 130-p but extending parallel thereto. Each metal bit line such as bit line 113-m contacts underlying drain regions such as drain region 111-i through a via and contact region such as 112-i. As explained above, drain region 111-i and comparable drain regions 211-i through N11-i and 115-i through N15-i each serve as the drain regions for two self-aligned transistors such as transistors 110-i and 110-(i+1). Each source line 142 serves as the source for each of the 2N or 4N transistors connected thereto. However, each incremental section of the source line between a given pair of adjacent transistors (such as transistors 110-i and 210-i) has an incremental resistance R associated there-

with. Thus, when a read current I_r passes through each incremental resistance there is a drop in voltage by the amount I_rR with the result that the drain to source voltage difference at transistor N10-i is reduced by the amount NI_rR. Accordingly, the voltage drop generated by the read current passing through the source regions to metal contact line 130-p places a practical limit on the maximum size of the number N.

As shown schematically in FIG. 9*c* source region 142-s serves as the source not only for transistors 110-i+1 through N10-(i+1) and 114-(i+1) through N14-(i+1), but also for transistors 110-(i+2), 210-(i+2) through N10-(i+2) and transistors 114-(i+2) through N14-(i+2). Thus the second through subsequent rows of source lines 142 each are connected to two lines of transistors.

The metal bit lines 113-m through N13-m and 113-(m+1) through N13-(m+1) are each connected in rows to a plurality of pairs of transistor cells in the same manner as shown in conjunction with line 113-m, the contact region 112-i and transistor cells 110-i and 110-(i+1) (FIGS. 9*a* and 9*b*). Typically a row of the array will include 256, 512 or more metal contact regions such as contact region 112-i contacting the drains of 512, 1024 or more transistor cells such as cell 110-i and cell 110-(i+1).

Among the advantages of the structure shown in FIGS. 9*a*, 9*b* and 9*c* is that the capacitance of a typical word line 142 to a typical bit line 113 is substantially reduced. The reason for this can be seen in FIG. 9*b*. In the prior art each time a word line 122 passes over a bit line 113, in the prior art a fairly high capacitance exists because the bit line has a high donor concentration (typically n+) and the oxide between the word line and the bit line is thin. However, with this invention, the word line 122-n is separated from the overlying bit line 113 by fairly thick oxide 150*a* (typically around one micron thick) and thus the capacitance is very low. Moreover, the word line 122-n has very little overlap associated with edge 144*a* overlying the drain region 111-i. Accordingly, the bit line to word line capacitance is substantially reduced.

Another advantage of this invention lies in the fact that the read current I_r is quite high because of the use of a split gate. With a split gate the floating gate 121-i can be made shorter but wider because high coupling to the gate is not required. Since the current is proportional to the width of the floating gate over the length of the floating gate, the current is increased by having a short floating gate Finally, the bit line capacitance itself is smaller than in the prior art because the bit line 113-m contacts 256 discrete drains rather than a continuous drain line. The reduction in size of the drain regions by using unconnected drain regions such as drain 111-i substantially reduces the capacitance associated with the bit line.

While several embodiments of this invention have been described, other embodiments of this invention will be obvious to those skilled in the semiconductor arts in view of this disclosure.

What is claimed is:

1. An EPROM array comprising:

a substrate composed of semiconductor material;

a plurality of memory cells formed on the substrate, each memory cell including a split gate transistor;

a metal source contact line running in a first direction across the array;

a source diffusion line having a multiplicity of portions serving as source regions of the split gate transistors, the source diffusion line being integrally formed in the substrate and running a selected distance across said array orthogonal to said metal source contact line;

a contact between said metal source contact line and the source diffusion line for coupling a potential on said metal source contact line to each of said source regions;

a plurality of metal drain lines running across said array substantially parallel to said metal source contact line each metal drain line contacting drain region of a selected number of the split gate transistors in said array;

a plurality of control lines formed over said array running orthogonal to said metal source contact line and said plurality of metal drain lines;

wherein each split gate transistor comprises: a channel region; a floating gate formed over but insulated from a first portion of the channel region, a first edge of the floating gate being aligned with and used to define one edge of the drain region of the split gate transistor, a second edge of said floating gate being over said channel region, the second edge being positioned away from the first edge of the floating gate by a predetermined distance and separated from the closest edge of the source region of the transistor by a second portion of said channel region, and a control gate formed over but insulated from said floating gate and formed over but insulated from said second portion of said channel region; and

wherein said control gate of each split gate transistor comprises part of one of said plurality of control lines.

2. An EPROM array as in claim 1 wherein the source diffusion line has no more than 2N portions serving as source regions and N is an integer selected from the group consisting of 2, 4 and 8.

3. A memory device comprising the structure of claim 1 replicated M times to form an EPROM device with M times the memory cells of the structure of claim 1.

4. Structure as in claim 1 wherein said control gate extends over the second portion of said channel region to the source region and the edge of said control gate furthest from said floating gate is used to define, and is aligned with, the edge of the source region adjacent the channel region.

5. An EPROM array containing a plurality of memory cells wherein each cell in the array includes a transistor containing a source region, a drain region and a channel region therebetween, a floating gate formed over but insulated from a first portion of the channel region, a first edge of the floating gate being aligned with and used to define one edge of said drain region and a second edge of said floating gate being over said channel region, separated from the first edge of the floating gate by a predetermined distance, and separated from the closest edge of said source region by a second portion of said channel region, and a control gate formed over but insulated from said floating gate and formed over but insulated from said second portion of said channel region.

6. Structure as in claim 5 wherein said control gate extends over the second portion of said channel region to the source region and the edge of said control gate

furthest from said floating gate is used to define, and is aligned with the edge of the source region adjacent the second portion of the channel region and wherein the control gate further extends beyond the first edge of the floating gate.

7. A memory array comprising:

a semiconductor substrate; and

a plurality of split gate transistors formed in the substrate, each having: a source region; a drain region spaced apart from the source region; a first channel portion interposed between the source and drain regions; a second channel portion interposed between the first channel portion and the source region; a floating gate insulatively disposed over the first channel portion, the floating gate having opposed first and second edges spaced apart by a predetermined distance, the first edge of the floating gate being self-aligned with and used to define an edge of the drain region; and a control gate overlapping the second channel portion and the floating gate, the control gate having a portion extending over and beyond the first edge of the floating gate.

8. The memory array of claim 7 wherein the control gate of each transistor has an edge self-aligned with and used to define an edge of the source region.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 4,868,629

DATED : September 19, 1989

INVENTOR(S) : Boaz Eitan

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Sheets 1-7 of the drawing should be deleted to be replaced with sheets 1-7 of the drawing as shown on the attached sheets.

Signed and Sealed this

Thirteenth Day of March, 1990

Attest:

JEFFREY M. SAMUELS

Attesting Officer

Acting Commissioner of Patents and Trademarks

[redacted]

FIG. 1
PRIOR ART

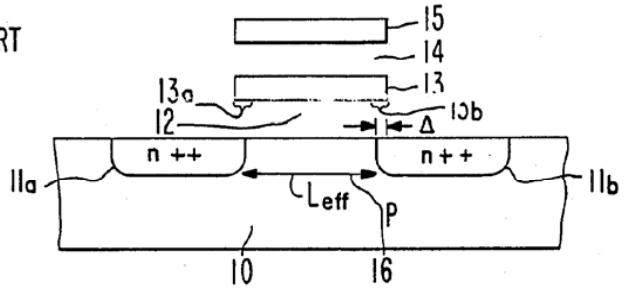


FIG. 2
PRIOR ART

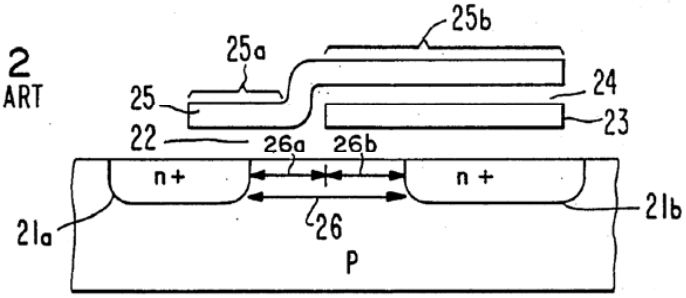


FIG. 3

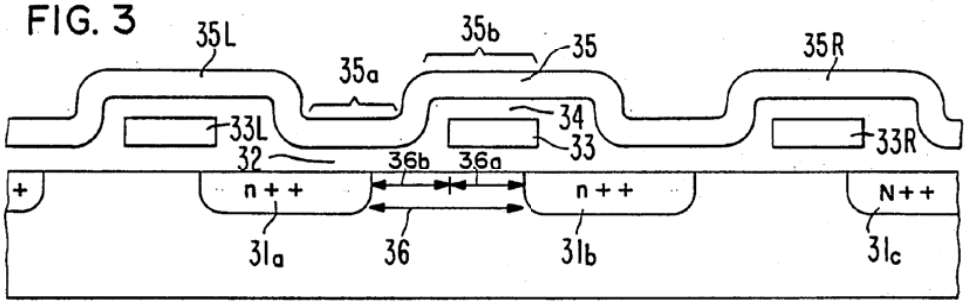
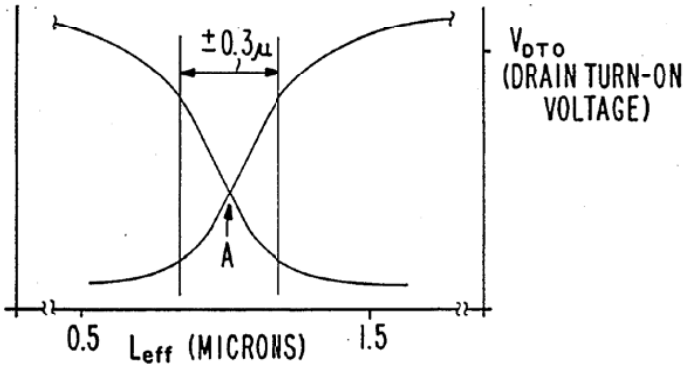


FIG. 4
PRIOR ART

V_{tx}
(THRESHOLD VOLTAGE)



[redacted]

KRNS000919

FIG. 5a

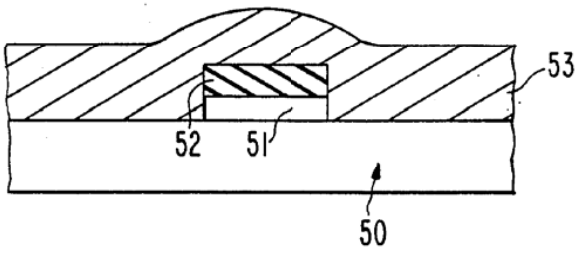


FIG. 5b

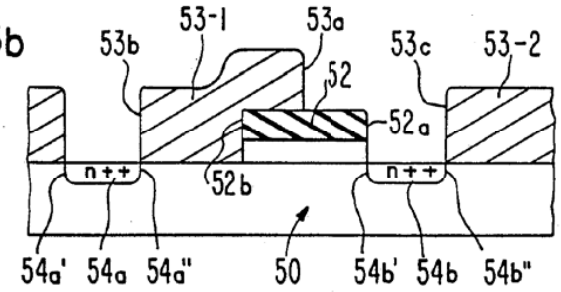


FIG. 6a

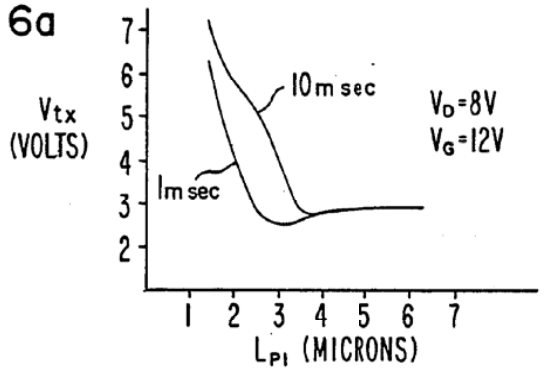
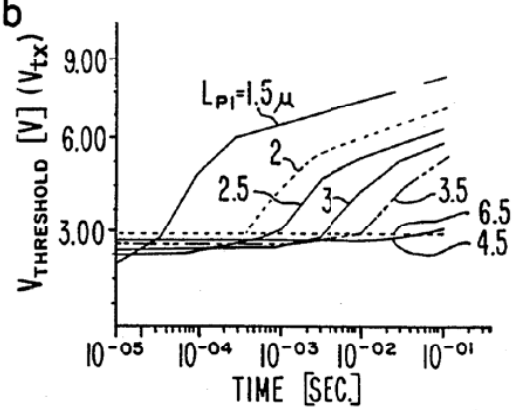


FIG. 6b



[redacted]

FIG. 6c

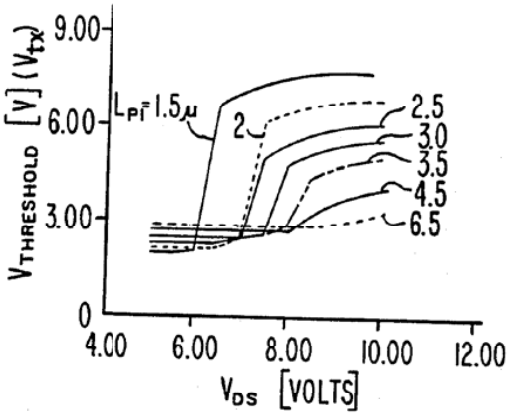
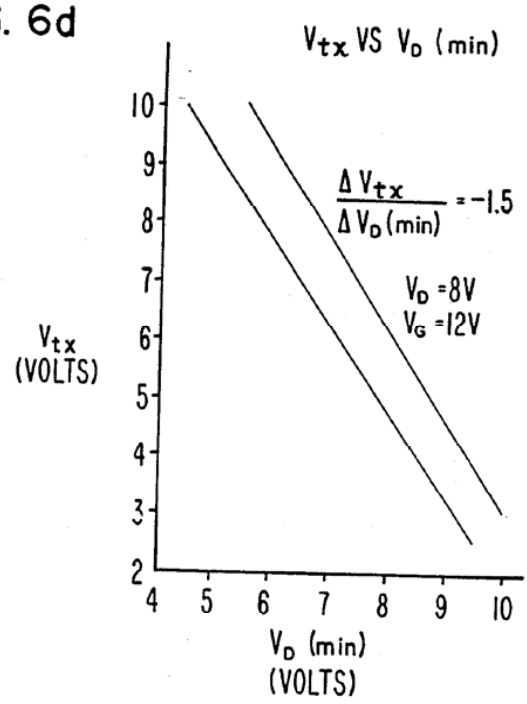


FIG. 6d



[redacted]

KRNS000921

FIG. 7a

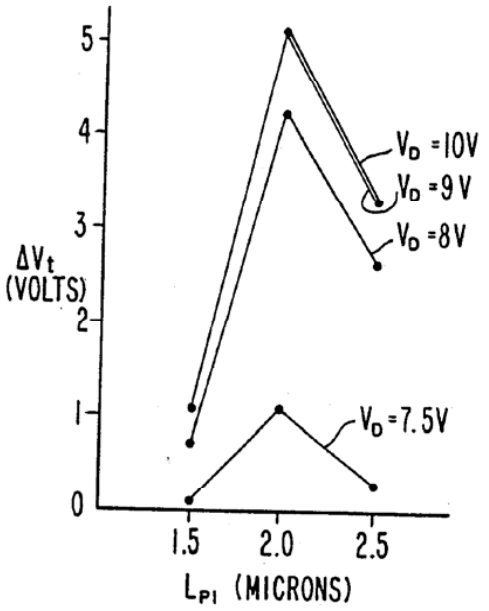


FIG. 7b

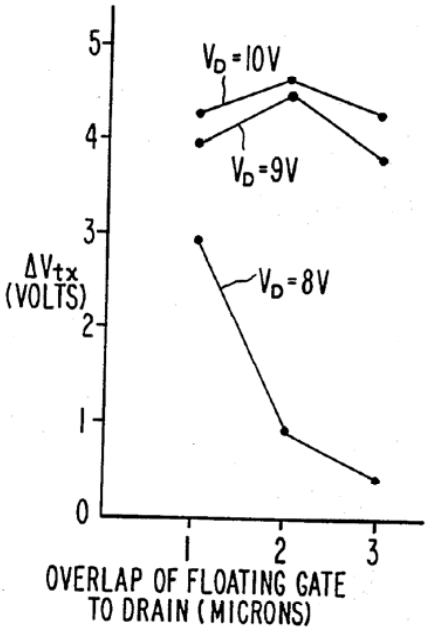
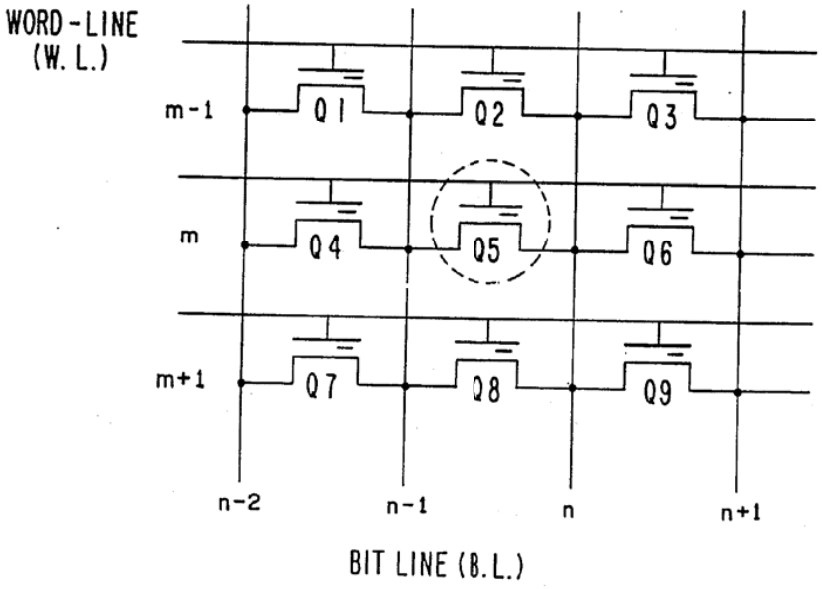


FIG. 8



[redacted]

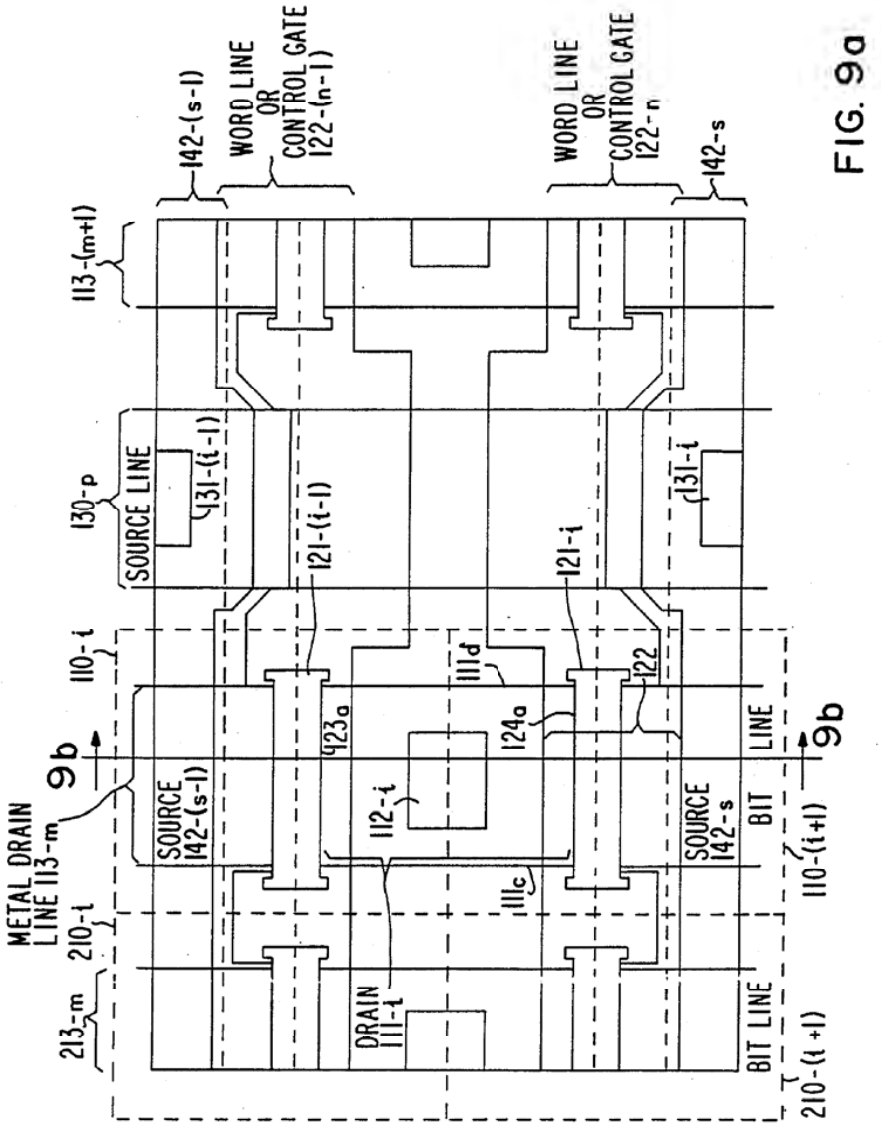


FIG. 9a

[redacted]

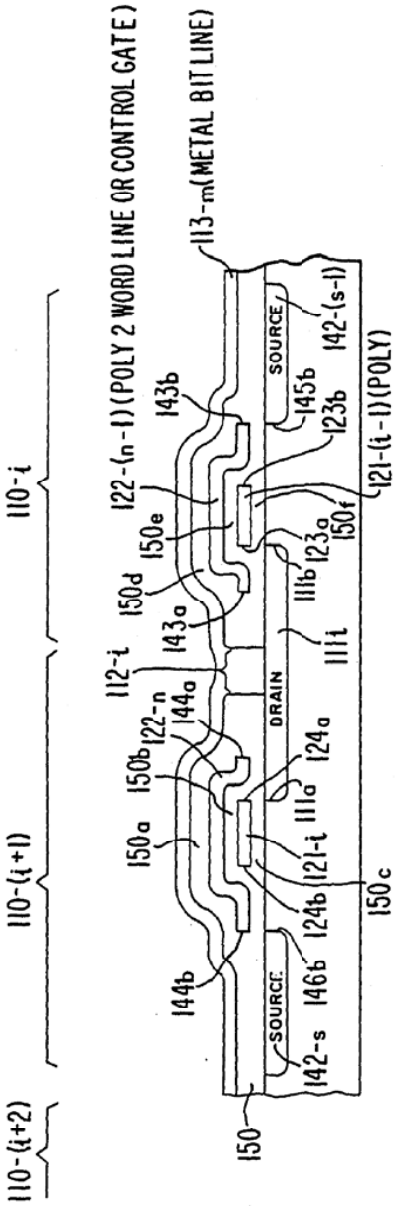


FIG. 9b

[redacted]

KRNS000925

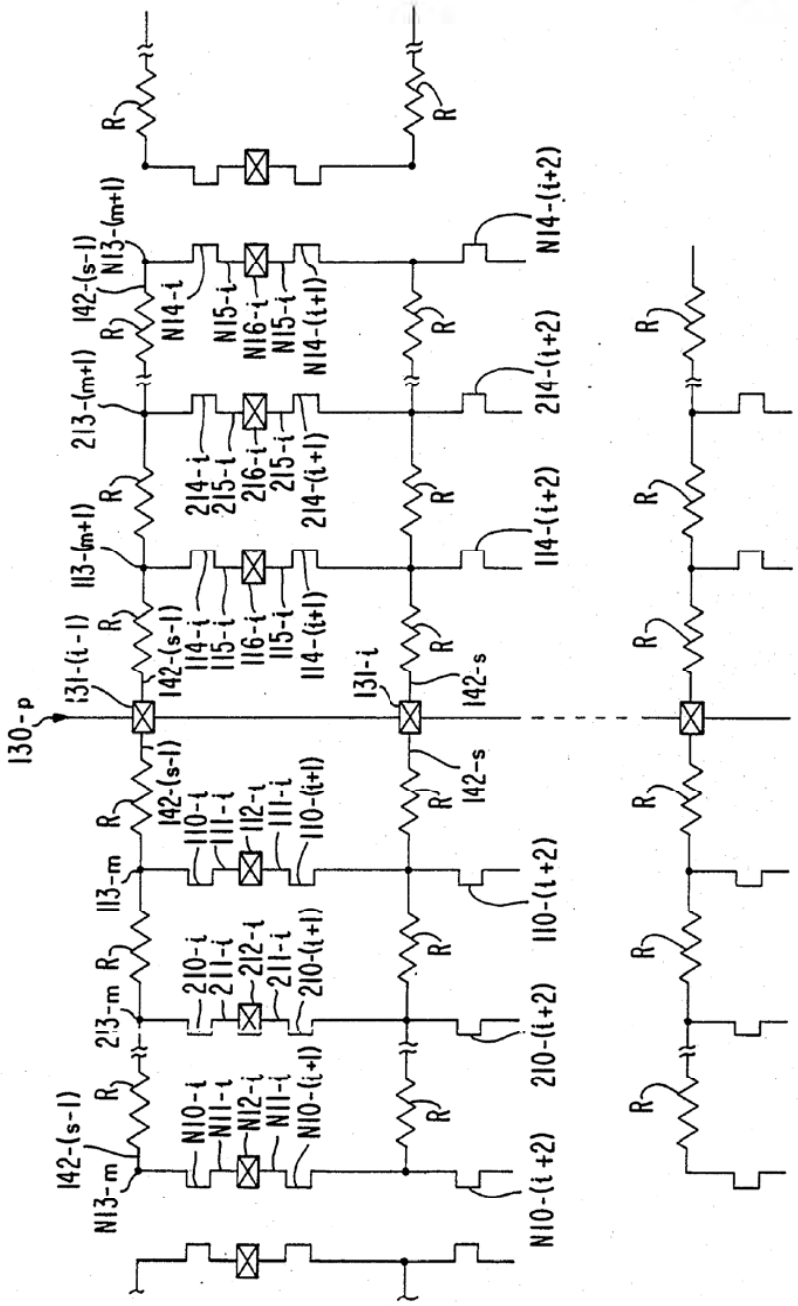


FIG. 9c

[redacted]

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 4,868,629 Page 1 of 3

DATED : September 19, 1989

INVENTOR(S) : Boaz Eitan

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Col. 1, line 21, insert --structure-- after "gate".

Col. 1, line 22, delete "1B".

Col. 1, line 27, delete "23".

Col. 1, line 50, delete "B".

Col. 1, line 56, insert --++-- after "n".

Col. 1, line 66, insert --.-- after "13".

Col. 2, line 9, "drain 11b" should read --and the drain 11b-- with number "11b" in bold.

Col. 2, line 22, "B" should be deleted.

Col. 5, line 26, the line should read --to form a floating gate 52. The oxide--.

Col. 5, line 54, change "n." to --n++--.

Col. 5, line 59, insert --segment-- after "photoresist".

Col. 7, line 9, "Leff should read --L_{eff}--.

Col. 7, line 21, insert --,-- after "rapidly".

Col. 7, line 23, insert --,-- after "increases".

Col. 7, line 49, "VDS" should read --V_{DS}--.

[redacted]

KRNS000926

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 4,868,629

DATED : September 19, 1989

INVENTOR(S) : Boaz Eitan

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Col. 8, line 8, insert --,-- after "reached".

Col. 8, line 24, "V_D" should read --V_{DS}--.

Col. 8, line 31, "print" should read --point--.

Col. 8, line 55, "ΔV_T" should read --ΔV_{tx}--.

Col. 9, line 12, "Three effective ..." should begin a new paragraph.

Col. 9, line 14, "ΔV_T" should read --ΔV_{tx}--.

Col. 9, line 24, "Vtx" should read --V_{tx}--.

Col. 9, line 33, delete "." after "microns".

Col. 9, line 57, "ΔV_T" should read --ΔV_{tx}--.

Col. 10, Table 1, insert --structure-- after "gate" in left column heading of table;
delete "structure" after "gate" in left column step 2, line 3, of table.

Col. 11, line 47, "+0.3" should read --$\pm$0.3--.

Col. 11, line 48, "+0.6" should read --$\pm$0.6--.

Col. 11, line 50, "L" should read --L_{Pl}--.

Col. 11, line 50, "+0.6 or +0.7" should read --$\pm$0.6 or $\pm$0.7--.

[redacted]

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 4,868,629 Page 3 of 3

DATED : September 19, 1989

INVENTOR(S) : Boaz Eitan

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Col. 11, line 54, "+0.3" should read --±0.3--.

Col. 12, line 11, insert --a-- after "is also in".

Col. 12, line 61, "121S3 (i-1)" should read --121-(i-1)--

Col. 15, line 50, insert "." after "gate".

Signed and Sealed this

Fourth Day of June, 1991

Attest:

Attesting Officer

HARRY F. MANBECK, JR.

Commissioner of Patents and Trademarks

[redacted]

KRNS000928

United States Patent [19]

Kazerounian et al.

[11] **Patent Number: 5,042,009**

[45] **Date of Patent: Aug. 20, 1991**

[54] **METHOD FOR PROGRAMMING A FLOATING GATE MEMORY DEVICE**

[75] Inventors: **Reza Kazerounian**, Alameda; **Boaz Eitan**, Sunnyvale, both of Calif.

[73] Assignee: **WaferScale Integration, Inc.**, Fremont, Calif.

[21] Appl. No.: **282,788**

[22] Filed: **Dec. 9, 1988**

[51] **Int. Cl.**[5] **G11C 13/00**

[52] **U.S. Cl.** **365/185**; 365/189.01; 365/230.01

[58] **Field of Search** 365/185, 189.01, 230.01

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,580,067	4/1986	Proebsting et al.	307/296 R
4,628,487	12/1986	Smayling	365/185
4,667,312	5/1987	Doung et al.	365/189
4,723,225	2/1988	Kaszubinski et al.	365/185
4,763,299	8/1989	Hazane	365/51

OTHER PUBLICATIONS

R. Kazerounian, et al., "A 5 Volt Only High Density Poly-Poly Erase Flash Eprom Cell", IEDM. Dec. 1988, pp. 436–439.

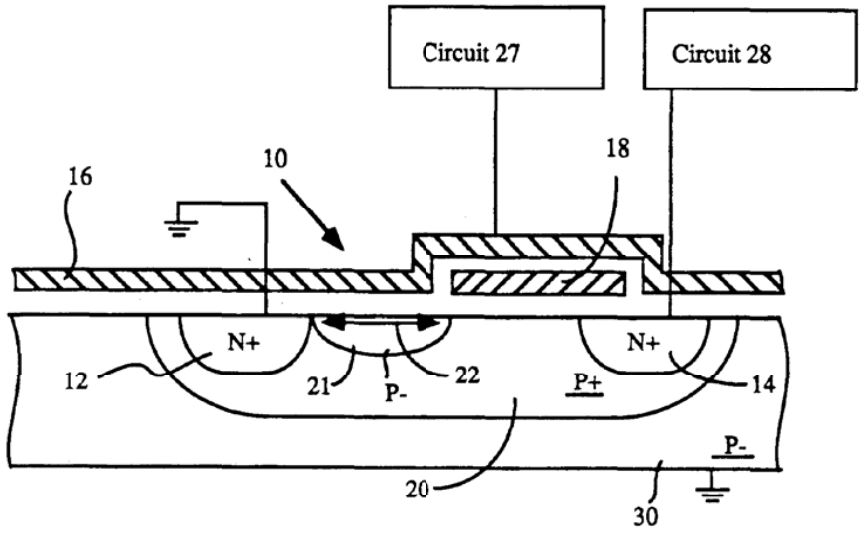
Primary Examiner—Terrell W. Fears

Attorney, Agent, or Firm—Skjerven, Morrill, MacPherson, Franklin & Friel

[57] **ABSTRACT**

A method of programming a floating gate transistor permits the use of a charge pump to provide drain programming current. The programming drain current is typically held below about 1 μA. This programming drain current can be provided by a conventional charge pump. In the first embodiment, the drain current can be limited by connecting a resistor between the source and ground. In a second embodiment, the drain current is limited by limiting the transistor control gate voltage. In a third embodiment, a charge pump is coupled to the drain while the control gate is repetitively pulsed. Each time the control gate is pulsed, the transistor turns on, and although the drain is initially discharged through the transistor, some hot electrons are accelerated onto the floating gate, and eventually the floating gate is programmed. In these embodiments the erase gate voltage may be raised to enhance programming efficiency.

30 Claims, 10 Drawing Sheets



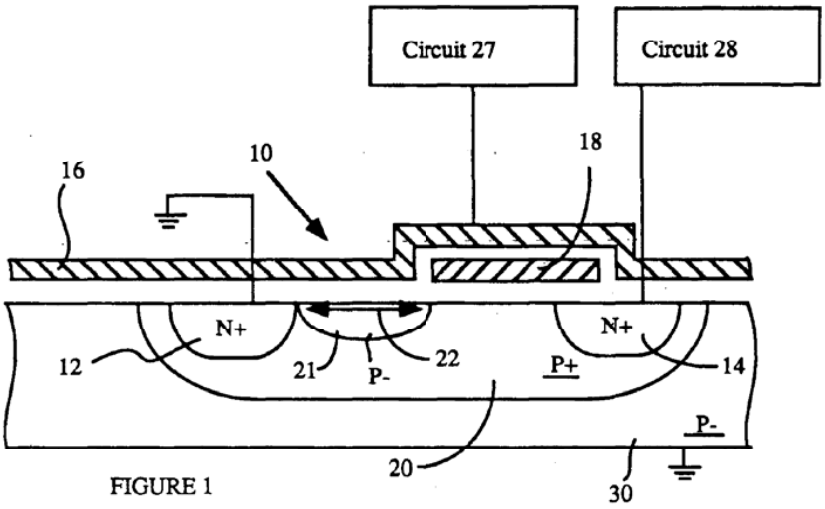


FIGURE 1

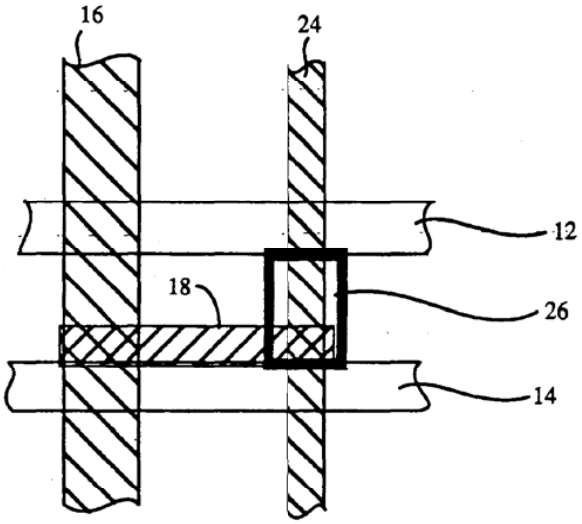
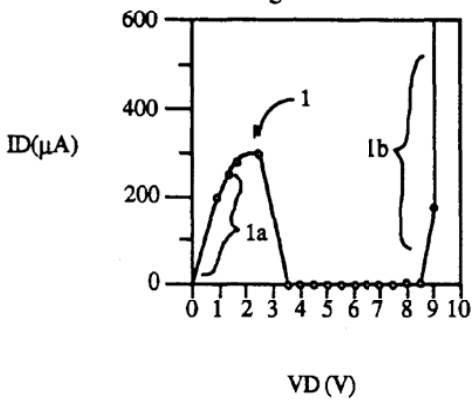


FIGURE 1a

Figure 2


600
400
200
0
ID(μA)
1
1a
1b
0 1 2 3 4 5 6 7 8 9 10
VD (V)

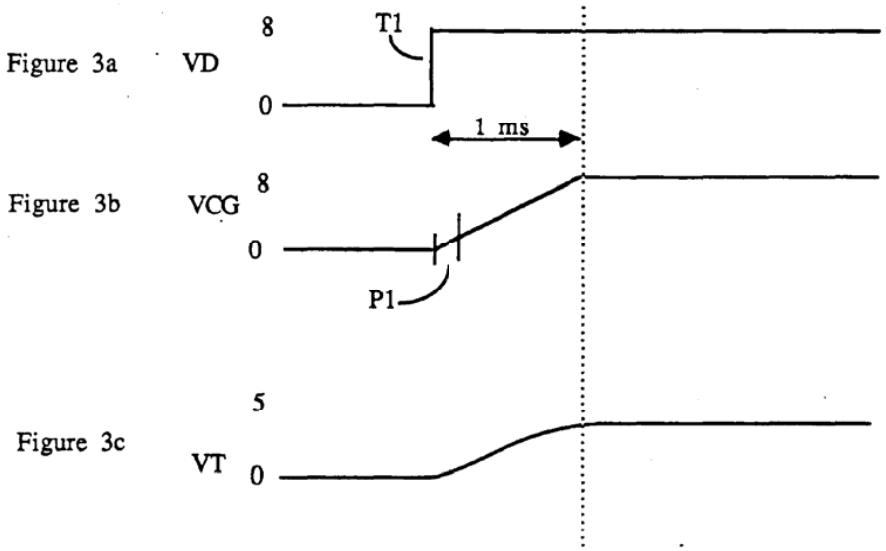


Figure 3a
VD
8
0
T1
1 ms
Figure 3b
VCG
8
0
P1
Figure 3c
VT
5
0

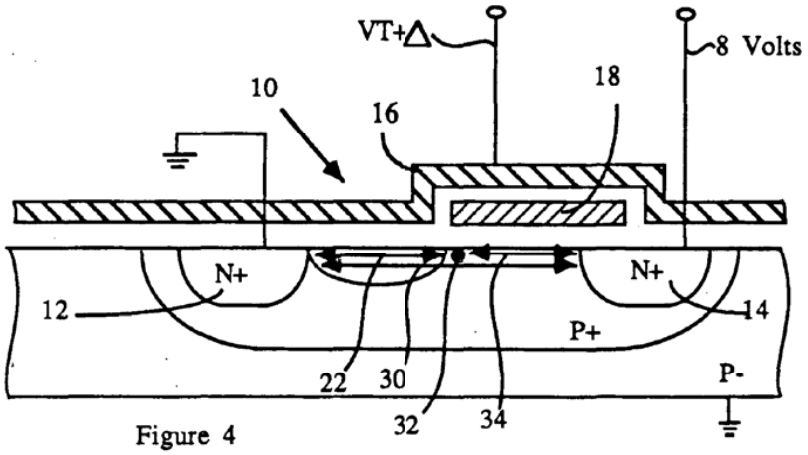


Figure 4

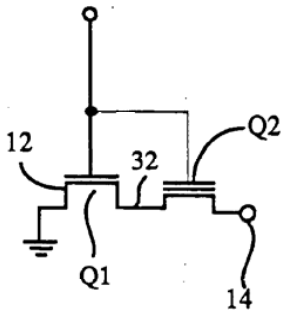


Figure 4a

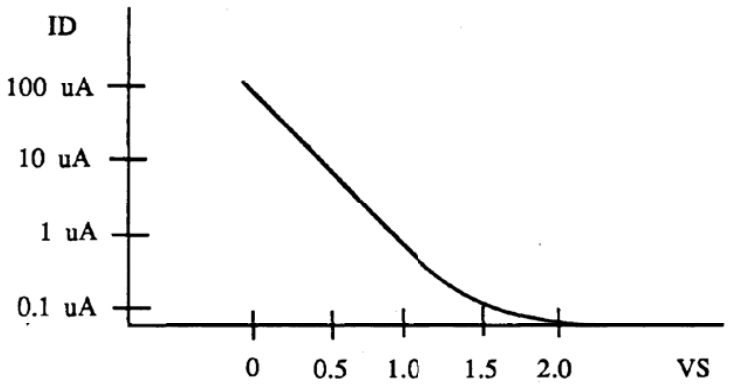


Figure 5

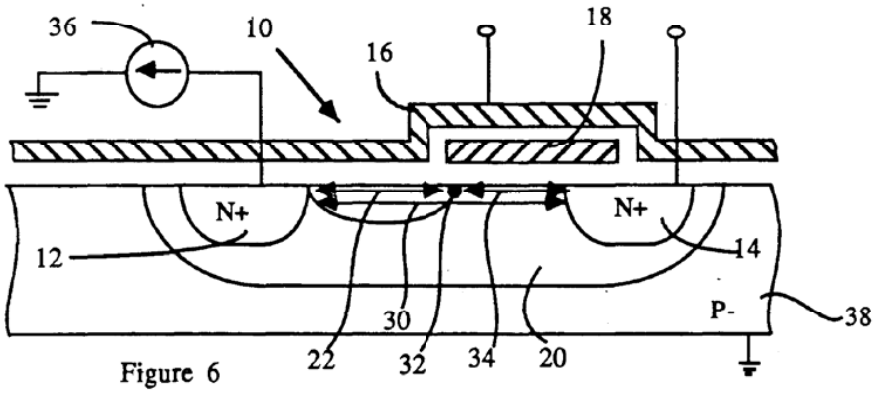


Figure 6

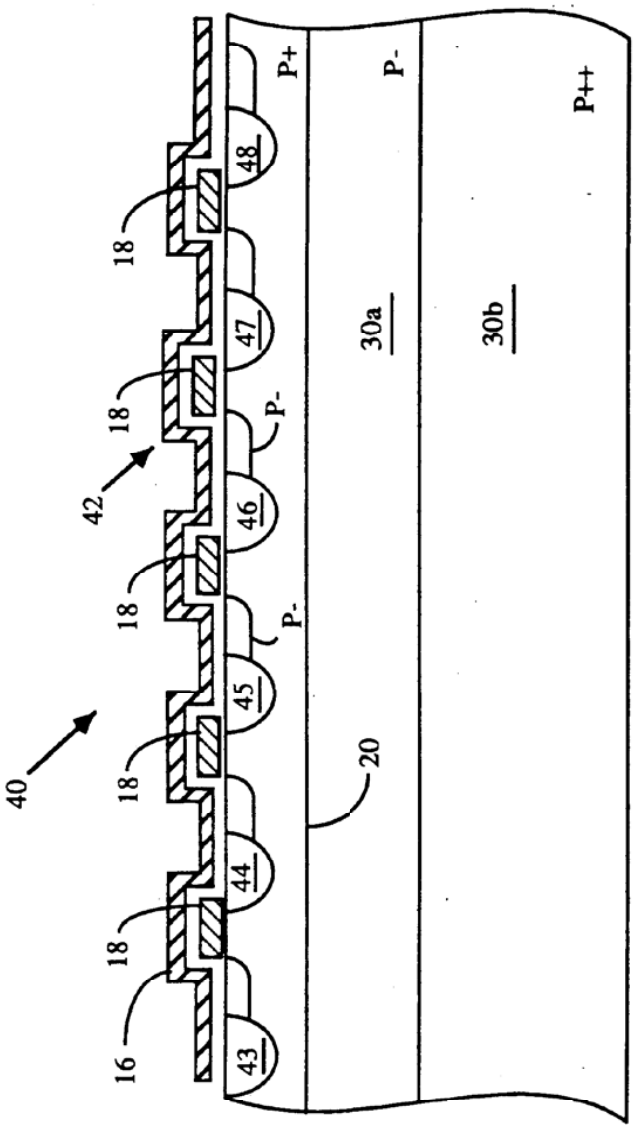


Figure 6a

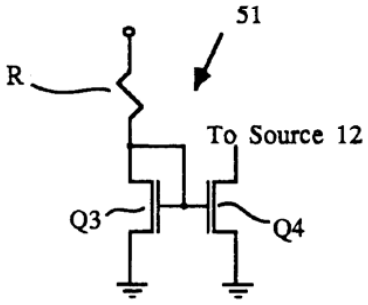


Figure 7

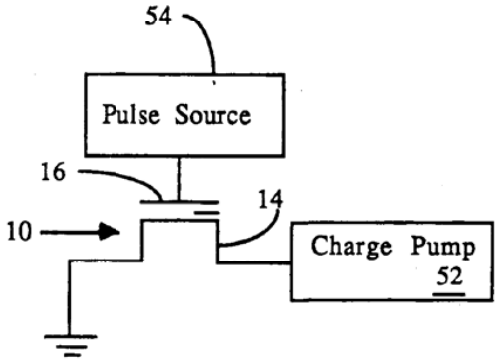


Figure 8

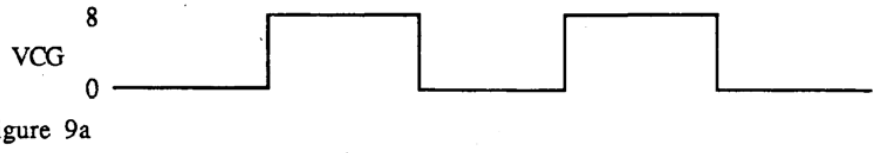


Figure 9a

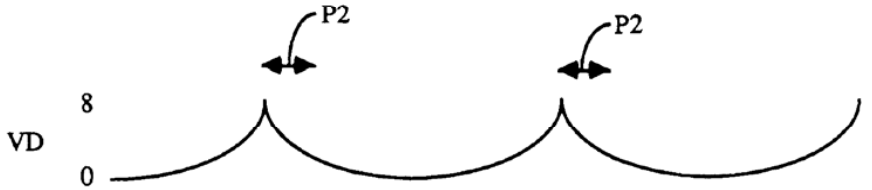


Figure 9b

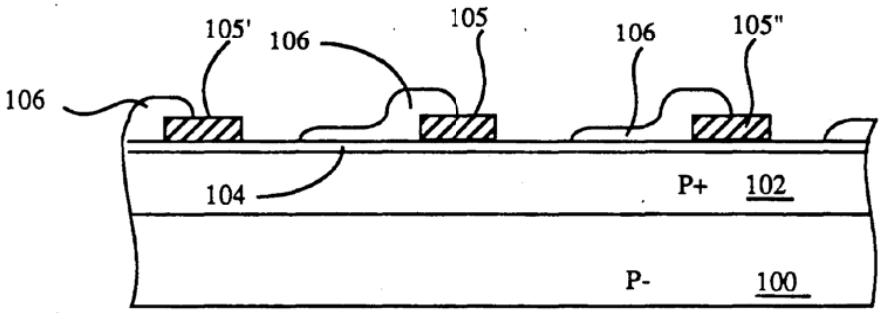


Figure 10a

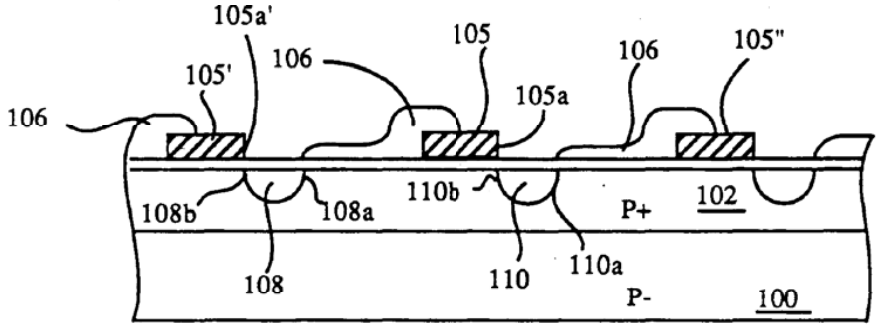


Figure 10b

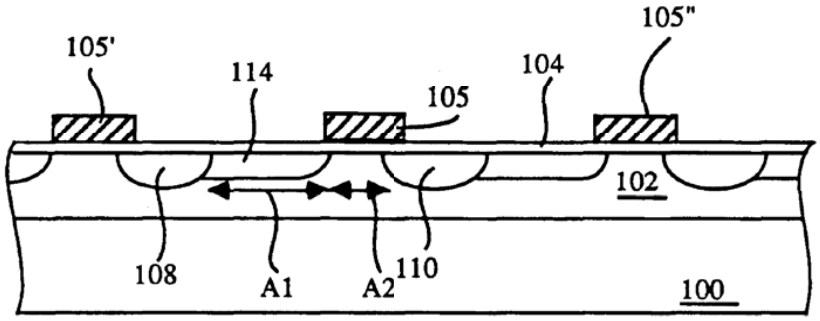


Figure 10c

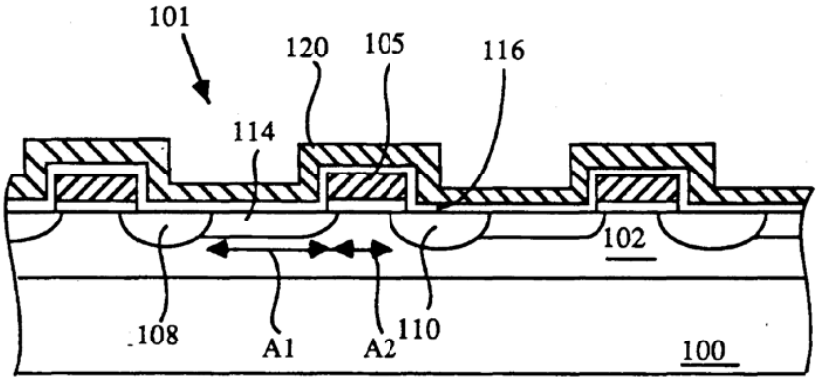


Figure 10d

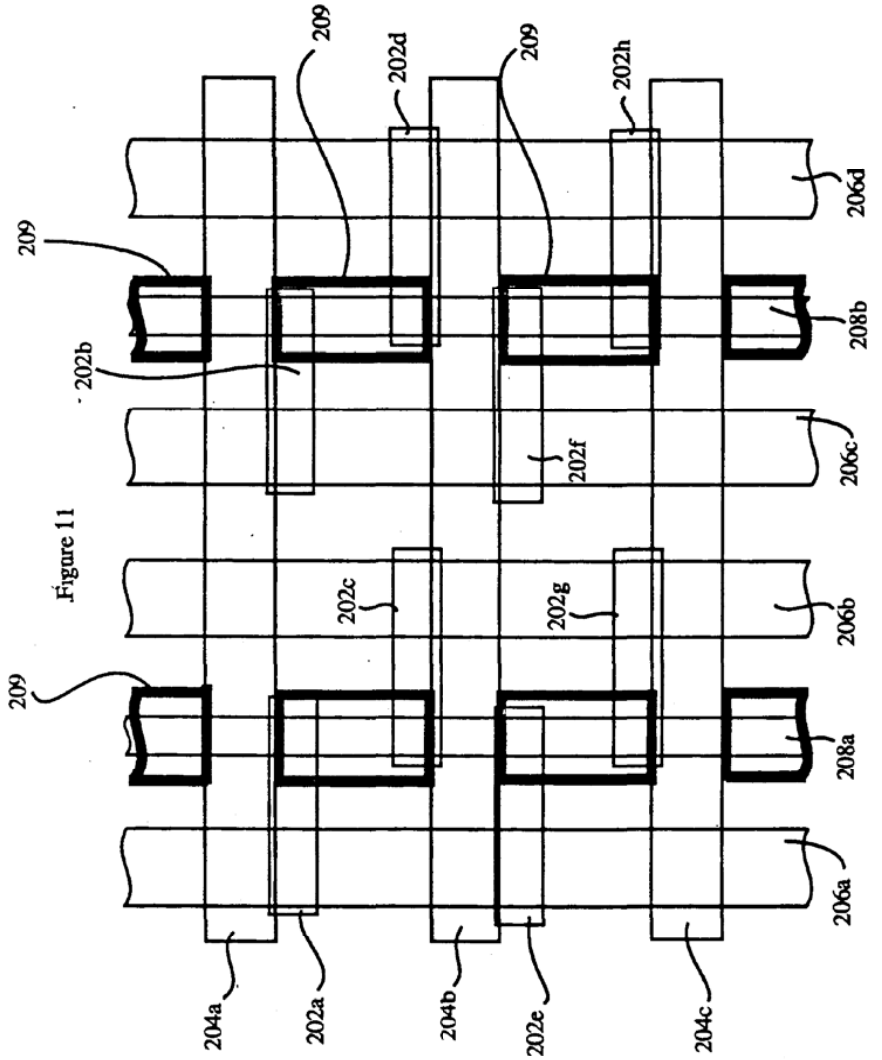


Figure 11

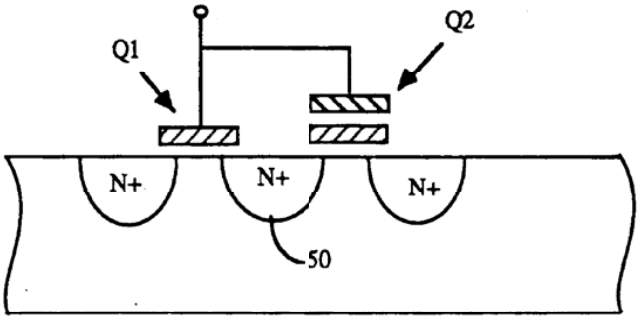


Figure 12

METHOD FOR PROGRAMMING A FLOATING GATE MEMORY DEVICE

BACKGROUND OF THE INVENTION

This invention relates to floating gate memory devices and methods for programming such devices.

There are a number of floating gate memory devices known in the art. One type of floating gate memory comprises an array of floating gate transistors which are programmed and erased by an electron tunneling mechanism. An example of such a device is discussed by Johnson et al. in "A 16 Kb Electrically Erasable Nonvolatile Memory," published at the IEEE International Solid State Circuits Conference in 1980, page 152–153, incorporated herein by reference. Johnson's device uses programming and erase voltages of about 25 volts. Although most digital electronic systems include a 5 volt power supply but do not include a 25 volt power supply, 25 volts can be generated on-chip from a 5 volt power supply with a conventional charge pump, since the amount of current required for tunneling is on the order of 1 nA. Unfortunately, memory cells which are programmed and erased by tunneling tend to be large, and thus expensive.

Another type of floating gate memory is the EPROM, which is programmed by hot electron injection and erased by exposure to UV light. EPROM cells are small, and are less expensive to build than EEPROM cells, but the data stored in the EPROM cannot be reprogrammed unless the EPROM is removed from a system and exposed to UV light prior to reprogramming. Further, such devices are programmed by hot electron injection, which requires a voltage in excess of 5 volts (e.g. about 12 volts) and a high programming current. Such programming currents are too large to generate with a charge pump. Thus, if one wanted to program an EPROM in-system, one would have to include an extra power supply, which would entail an undesirable expense.

Another type of floating gate memory is the flash EPROM, which is programmed by hot electron injection and erased by tunneling. Such a device is discussed by Kynett et al. in "An In-System Reprogrammable 256K CMOS Flash Memory", published at the IEEE International Solid State Circuits Conference in 1988, pages 132 to 133, incorporated herein by reference. Advantageously, flash EPROMs have small memory cells, and are thus relatively inexpensive. However, since flash EPROMs of the type discussed by Kynett are erased by electron tunneling either between the floating gate and drain or between the floating gate and source, they draw a large current during electrical erase due to band to band tunneling across the drain/substrate or source/substrate junction. Flash EPROMs also have a number of other disadvantages. For example, they are hot electron programmed, and thus require a programming voltage in excess of 5 volts (typically 8 to 12 volts) with about 1 mA of programming current per cell. This combination of high current and high programming voltage cannot be economically generated from an on-chip charge pump. (Flash EPROMs cannot be efficiently programmed merely by connecting a 5 volt power supply to the drain, especially at high operating temperatures, e.g. 125° C. Also, since the output voltage of a nominally 5 volt power supply may vary by plus or minus 10%, and thus be as low as 4.5 volts, programming cannot be efficiently accomplished by connecting the 5 volt power supply to the drain for this reason as well.) Another limitation of the above flash EPROM is the need for a tightly regulated erase voltage to prevent over-erase, i.e. to prevent the erase circuitry from leaving the floating gate with a large positive charge. (Since Kynett's floating gate extends from the source to the drain, a positively charged floating gate would leave Kynett's transistor on regardless of the state of his control gate.)

It would be desirable to provide a floating gate memory device which combines the following features:

(1) The small cell size of a flash EPROM;

(2) The erasability of an EEPROM, i.e. a device which can be erased in-system, wherein the erase voltage is generated by a charge pump from a single 5 volt power supply; and

(3) In-system programmability from a single 5 volt power supply.

These goals could be achieved if a method were found for programming a flash EPROM without requiring more than a few microamps of drain current.

SUMMARY

A erasable floating gate memory device constructed in accordance with an embodiment of the invention has the small cell size of a flash EPROM, but can be programmed and erased using a single 5 volt power supply. Of importance, the programming and erase voltages are generated on-chip from the 5 volt power supply, e.g. using a charge pump.

One embodiment of the invention includes means for limiting the amount of current permitted to flow through the drain during programming. Because of this, the programming drain voltage can be generated by a charge pump and it is not necessary to provide an additional power supply for programming the memory device.

In a first embodiment, during programming, the control gate voltage of the floating gate memory device is ramped from a first voltage (e.g. ground) to a programming voltage (e.g. between 5 and 8 volts) over a time period such as 1 millisecond. Because of this, the programming drain current ramps up slowly during the 1 millisecond period, hot electrons are continuously injected onto the floating gate during the 1 millisecond period, the threshold voltage of the transistor is constantly increasing, and there is no period of time during which the drain current exceeds a value greater than that which the charge pump can provide.

In a second embodiment of the invention, during programming, the control gate is raised to a value just slightly greater than the threshold voltage of the transistor while a programming drain voltage is applied to the drain region. This ensures that the drain current through the transistor is of a magnitude which can be provided by a charge pump. The memory device typically incorporates a split gate architecture, i.e. the floating gate covers a first portion of the channel but not a second portion. The control gate covers the second portion of the channel and part of the floating gate. Thus, the control gate controls the amount of current permitted to flow through the channel, even if the floating gate is positively charged.

In this embodiment, the memory device includes an erase gate which is capacitively coupled to the floating gate. During programming, the erase gate voltage is raised, e.g. to about 10 volts, to thereby increase the

electrical potential at the floating gate in order to enhance the programming efficiency of the memory device. It is thus seen that the control gate is used to control the amount of programming current, while the erase gate enhances programming efficiency.

In another embodiment of the invention, during programming, the source is coupled to ground via a current limiting element. The current limiting element limits the source current to a value between 1 and 5 μA. An example of such an element is a 1 megaohm resistor. This raises the source voltage during programming, thereby increasing the threshold voltage of the transistor due to the back bias effect, thus reducing the amount of drain current permitted to flow during programming. Because of this increase in threshold voltage, the programming current permitted to flow between the source and drain is limited to a value which can be generated by the charge pump. In this embodiment, in addition to coupling a current limiting element between the source and ground, the transistor erase gate voltage is raised, e.g. to about 10 volts. Since the floating gate is capacitively coupled to the erase gate, this has the effect of increasing the floating gate voltage and enhancing programming efficiency. However, in other embodiments, the erase gate is grounded during programming.

In yet another embodiment, the charge pump is coupled to the transistor drain region while the control gate is periodically pulsed. When the control gate voltage is low, the drain voltage rises to about 8 volts. When the control gate is pulsed, the drain is discharged through the floating gate transistor, and when the control gate voltage is low again, the drain region is permitted to charge to 8 volts. As described in greater detail below, repeatedly pulsing the control gate permits one to program the floating gate transistor with a charge pump, even though the charge pump cannot provide more than a few microamps of current. This programming technique can be used while raising the erase gate voltage to enhance programming efficiency or in conjunction with a grounded erase gate.

In one embodiment, the memory device comprises a staggered virtual ground array of split gate floating gate memory cells. The array comprises a set of elongated source/drain regions and a plurality of rows of floating gates, each row of floating gates formed between a pair of source/drain regions. The floating gates are arranged so that in a given row, every other floating gate is adjacent to a first one of the source/drain regions within the pair, and the remaining floating gates within the row are adjacent a second one of the source/drain regions within the pair. Because of this, the array can be constructed in a smaller surface area than would be possible if all of the floating gates in a given row were adjacent the same source/drain region.

In one embodiment, each cell comprises a channel region between a pair of associated source/drain regions. The channel region includes a first portion under the floating gate (and adjacent to one of the source/drain regions within the pair) which is heavily doped, and a second portion adjacent the other source/drain region which is more lightly doped. The first portion of the channel enhances the programming efficiency of the cell, while the low dopant concentration of the second portion of the channel causes the second portion of the channel to exhibit a low threshold voltage. We have discovered a novel method for doping the channel so that the first and second portions of the channel are self-aligned with the edges of the floating gate. This is done by (1) heavily doping the entire channel region, (2) forming the floating gate, and (3) partially counterdoping the portion of the channel that is not under the floating gate, using the floating gate as a mask. This technique improves yields because it is impossible to misalign the first and second portions of the channel with respect to the rest of the transistor.

These and other advantages of the present invention are better understood with reference to the detailed description below.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a circuit for programming a flash EPROM in accordance with a first embodiment of our invention.

FIG. **1***a* illustrates in plan view the flash EPROM of FIG. **1**.

FIG. **2** illustrates the drain current versus drain voltage characteristic curve of the transistor of FIG. **1** with a constant control gate voltage.

FIGS. **3***a* to **3***c* are waveform diagrams illustrating the drain voltage, control gate voltage and threshold voltage during programming of the transistor of FIG. **1**.

FIG. **4** illustrates a floating gate transistor which is programmed using a method in accordance with a second embodiment of our invention.

FIG. **4***a* schematically illustrates a circuit equivalent to that shown in FIG. **4**.

FIG. **5** illustrates the effect of a source bias voltage on drain current.

FIG. **6** illustrates a floating gate transistor in which the source is biased with respect to the substrate during programming by a current limiting circuit element.

FIG. **6***a* illustrates a row of transistors constructed in accordance with the embodiment of FIG. **6**.

FIG. **7** illustrates a circuit which provides an MOS transistor equivalent of a resistor which is used as the current limiting circuit element of FIG. **6**.

FIG. **8** illustrates a transistor, the drain of which is coupled directly to a charge pump and the control gate of which is coupled to a pulse source.

FIGS. **9***a* and **9***b* illustrate the control gate and drain voltage waveform applied to the floating gate transistor of FIG. **8**.

FIGS. **10***a* to **10***d* illustrate a floating gate transistor during a manufacturing process in accordance with our invention.

FIG. **11** illustrates in plan view a flash EPROM array constructed in accordance with our invention.

FIG. **12** illustrates a floating gate transistor coupled to a single gate transistor.

DETAILED DESCRIPTION

FIG. **1** illustrates in cross section a flash EPROM transistor **10** coupled to programming circuitry. Referring to FIG. **1**, transistor **10** includes an N+source **12**, an N+drain **14**, a control gate **16** and a floating gate **18**. Transistor **10** is formed within a P+region **20** to enhance the programming efficiency of transistor **10**. A P−region **21** is formed in a portion **22** of the channel region to reduce the effective threshold voltage of portion **22**. Transistor **10** also includes an erase gate **24** which extends over but is insulated from floating gate **18**. Erase gate **24** is outside of the cross section of FIG. **1**, but is illustrated in plan view in FIG. **1***a*. The portion of floating gate **18** which extends under erase gate **24** is formed over a field oxide region **26**.

If one were to attempt to program transistor **10** by raising the control gate and drain voltages to a programming voltage (e.g., about 8 to 12 volts), the drain current would initially rise to several hundred microamps as the drain voltage increased. The "one-shot" drain current versus drain voltage characteristic curve of transistor **10** is illustrated in FIG. **2**. Curve **1** in FIG. **2** indicates that current initially increases (portion **1***a*), but thereafter drops as floating gate **18** is programmed. If the drain voltage keeps increasing, current rises again (portion **1***b*) due to injection induced breakdown between drain **14** and P+region **20**. A charge pump cannot economically provide the several hundred microamps required to get past portion la of curve **1**. However, we have discovered a method for programming transistor **10** without providing such a large drain current.

In accordance with one embodiment of our invention, a control gate voltage waveform as illustrated in FIG. **3***b* is applied to control gate **16** by a circuit **27** while the voltage **12** waveform of FIG. **3***a* is applied to drain **14** by a circuit **28**. As can be seen, at or after a time Tl when a programming drain voltage of about 8 volts is applied to drain **14**, the voltage at control gate 16 ramps up from 0 to 8 volts over a time period (typically 0.1 to 10 ms, and preferably 1 ms). During a first portion P1 of this 1 ms period, control gate voltage VCG never exceeds a few volts, and the conditions required to draw a drain current of more than 1 μA never exist. However, during portion Pl, electrons are slowly injected onto floating gate **18**, and threshold voltage VT slowly starts to rise (FIG. **3***c*).

After portion Pl, control gate voltage VCG continues to increase to 8 volts. However, transistor **10** still does not draw more than 1 μA because threshold voltage VT also continues to increase, and conditions are never created which would permit a large drain current to flow. By the time voltage VCG reaches 8 volts, threshold voltage VT reaches about 8 volts, and transistor **10** is programmed without ever requiring more than 1 μA of drain current.

It will be apparent to those skilled in the art how to build circuits **27** and **28** capable of generating the voltage waveforms of FIGS. **3***a* and **3***b*. Thus circuits **27** and **28** will not be described in further detail herein except to note that the voltage applied to control gate **16** and drain **14** by circuits **27**, **28** is derived from a charge pump. Although 8 volts are applied to the transistor of FIG. **1**, this value is merely exemplary, and other voltages can also be applied to transistor **10**.

Transistor **10** is read in a conventional manner, e.g., by raising the voltage at control gate **16** to about 5 volts, raising the voltage at drain **14** to 1.5 volts, grounding source **12** and erase gate **24**, and sensing whether current flows through transistor **10**. Transistor **10** is erased by grounding control gate **16**, drain **14** and source **12** and raising the erase gate voltage to about **25** volts, thereby causing electrons to tunnel from floating gate **18** to erase gate **24**. This leaves floating gate **18** positively charged.

Although the embodiment discussed above functions adequately and comes within the scope of the invention, it does have some drawbacks. For example, different transistors in the array may be programmed at different rates. Assume, for example, that hot electrons reach the floating gate of one of the transistors in the array at a low rate. If the control gate voltage of that transistor increases too rapidly, the transistor will start to draw a large current (e.g. in excess of 100 μA) before charge is injected into its floating gate. Thus, the drain voltage V_D will start to drop, and programming will cease. Accordingly, the ramp rate must be selected to rise as slowly as the programming rate of the slowest transistor in the array allows. (If the ramp rate is too slow, programming will take too long.)

FIG. **4** illustrates another embodiment of my invention. In FIG. **4**, during programming a voltage of about 8 volts is applied to drain **14** while a voltage $V_T+\Delta$ is applied to control gate **16**, where V_T is the threshold voltage which, if applied to control gate **16**, will permit up to 1 μA to flow through portion **22** of transistor channel **30**. Δ is an incremental voltage, such that if VT +Δ is applied to control gate **16**, several microamps will be **34** permitted to flow through portion **22**. (V_T is typically about 1.0 volt, while Δ is about 0.2 volts.) It is thus seen that as long as the control gate voltage is less than or equal to $V_T+\Delta$, the drain current will be less than several microamps, and thus the transistor of FIG. **4** can be programmed using a conventional charge pump.

As is known in the art, the higher the electrical potential at floating gate **18**, the greater the programming efficiency of transistor **10**. In one embodiment, the electrical potential of floating gate **18** is enhanced by raising the voltage at erase gate **24**. Because of capacitive coupling between erase gate **24** and floating gate **18**, the increase in erase gate voltage, e.g. to about 10 volts, enhances programming of transistor **10**. Of course, the erase gate voltage cannot be raised too high, e.g., greater than 20 volts, or electrons will tunnel off of floating gate **18** and onto erase gate **24**.

It should be noted that although control gate **16** is biased such that portion **22** of channel **30** limits current below a few microamps, the voltage drop across portion **22** is only between **2** and **3** volts, even when **8** volts are applied to drain **14**. The reason for this is that transistor **10** can be envisioned as two transistors, i.e. a first transistor Q1 (FIG. **4***a*) whose source is source **12**, whose channel is channel portion **22**, and whose drain is point **32** between channel portions **22** and **34**. The drain, channel and source of the second transistor Q2 comprise drain **14**, channel portion **34**, and point **32**, respectively. As point **32** is biased with respect to P+region **20**, the back bias effect (also known as the body effect) of second transistor Q2 increases the effective threshold voltage of the second transistor, thus ensuring a large voltage drop between drain **14** and point **32**. (The relation between the source-substrate voltage and drain current for a transistor is illustrated in FIG. **5**). It is this voltage drop which accelerates hot electrons onto floating gate **18**. The enhanced dopant concentration at P+region **20** increases the back bias effect exhibited by second transistor Q2. (The back bias effect is discussed at pages 32 to 43 of "MOS Field-Effect Transistors and Integrated Circuits" by Paul Richman, published by John Wiley and Sons in 1973, incorporated herein by reference.) It is thus seen that the transistor of FIG. **4** is programmed without requiring a large drain current.

Although the embodiment of FIG. **4** functions adequately and comes within the scope of the present invention, it too has several drawbacks. For example, because dopant concentrations, oxide thicknesses and other parameters vary over the wafer surface area, the threshold voltages of the various transistors in the array may vary, and it may be difficult to generate a control gate voltage $V_T+\Delta$ which will permit programming of

the various flash EPROM transistors at an acceptable rate without permitting too much drain current to flow, and thus cause the drain voltage to drop.

In another embodiment, instead of applying about **8** volts to drain **14** and $V_T+\Delta$ to control gate **16**, a voltage of about **6** volts is applied to drain **14** and a voltage between about **2.5** and **3.5** volts is applied to control gate **16**. This will permit a programming drain current between about 50 μA and 100 μA. It is noted that while it is difficult to economically generate 100 μA from a charge pump which generates 8 volts from a 5 volt. 10% supply, 100 μA can be economically generated from a charge pump which generates 6 volts from a 5 volt±10% supply.

Since this embodiment permits between 50 and 100 μA to flow through the transistor, the voltage at point **32** will be lower in this embodiment than in the embodiment in which $V_T+\Delta$ is applied to control gate **16**. Since it is the voltage difference between drain **14** and point **32** which provides electrons with enough energy to reach floating gate **18**, the smaller drain voltage in this embodiment is offset by the lower voltage at point **32**.

It should be noted that in this embodiment, the control gate voltage need not be regulated as tightly as the embodiment in which $V_T+\Delta$ is applied to control gate **16**. Also, in this embodiment, the erase gate voltage is raised, e.g. to a voltage generally less than 15 volts and preferably about 10 volts.

FIG. **6** illustrates another embodiment of our invention in which programming drain current is held below 1 μA automatically without requiring the generation of a control gate voltage within very tight constraints. Referring to FIG. **6**, 8 volts are applied to drain **14**, about 4 volts are applied to control gate **16**, and a current limiter **36** is coupled between source **12** and ground during programming. (During reading and erasing, source **12** is connected directly to ground.) Current limiter **36** is typically a 1 MΩ resistor which limits the amount of current permitted to flow through transistor **10**. As current flows through transistor **10**, current limiter **36** has the effect of biasing source **12** relative to substrate **38** to generate the above-mentioned back bias effect by virtue of the ohmic voltage drop across the resistor. As the voltage at source **12** reaches about one volt, the back bias effect of first transistor Q1 causes the programming current to drop to about 1 μA. This causes the voltage at point **32** to rise, e.g. to a value between 2 and 3 volts, thereby increasing the back bias effect of transistor Q2. Because of this, transistor Q2 limits the drain current flowing through drain **14**, thereby ensuring that point **32** is at a voltage such that the voltage drop between drain **14** and point **32** is sufficient to accelerate hot electrons onto floating gate **18**. (Because of the enhanced dopant concentration of channel portion **34**, the drain current of transistor Q2 is more sensitive to its source voltage than transistor Q1. P+region **20** is grounded via its electrical connection to grounded substrate **38**.)

As electrons are accelerated onto floating gate **18**, the threshold voltage of transistor Q2 starts to increase, and the voltage at point **32** starts to decrease, so that the voltage across the source and drain of transistor Q2 increases. This increase in voltage facilitates further injection of hot electrons onto floating gate **18**.

As in the embodiment described above in relation to FIG. **4**, the erase gate voltage is typically raised during programming, e.g. to about 10 volts, to enhance programming efficiency.

After electrical erase, floating gate **18** is typically positively charged. This positive charge also effectively raises the floating gate electrical potential to further enhance programming efficiency.

As mentioned above, only about 4 volts are applied to control gate **16** during programming. The reason for this is that a transistor in accordance with the embodiment of FIG. **6** is typically part of a row of transistors such as row **40** of FIG. **6***a*. This row comprises a plurality of source/drain regions **43** to **48**, and the junction between each source/drain region and P+region **20** forms a capacitor. If it were desired to program a transistor **42** in row **40** and control gate **16** were raised to a voltage in excess of 6 volts, and all of floating gates **18** in row **40** were positively charged, all of source/drain regions **43** to **45** would be effectively connected to source/drain region **46** (source/drain **46** serves as the source of transistor **42**). That would be the equivalent of connecting a very large parasitic capacitance to source/drain region **46**, and it would take an unacceptably long amount of time to raise the voltage at region **46** and to program transistor **42**. By only raising the control gate voltage to only 4 volts, a resistance between source/drain region **46** and the other source/drain regions to the left of transistor **42** is created to reduce the effect of the above-mentioned parasitic capacitance. (In the embodiment of FIG. **6***a*, P− region **30***a* can be an epitaxial layer on a P++ substrate **30***b*.)

Transistor **10** of FIG. **6** is a split gate flash EPROM, meaning that floating gate **18** covers portion **34** of channel **30** but not portion **22**. As mentioned above, this is the equivalent of the pair of transistors Q1, Q2 in FIG. **4***a*. However, split gate transistor **10** has two advantages over an embodiment in which the EPROM cell was actually constructed as two transistors (FIG. **12**). First, the transistor of FIG. **6** (and FIGS. **1** and **4**) is smaller than transistors Q1 and Q2 of FIG. **12**. Second, in FIG. **6**, electrons gain energy while travelling from source **12** to point **32**, and for at least some of these electrons, this energy can be added to the energy gained by the electrons as they travel through channel portion **34**, to enhance programming efficiency. In the embodiment of FIG. **12**, any energy gained by electrons moving through the channel of transistor Q1 is completely lost as the electrons move through N+region **50**, and this lost energy cannot be used to enhance programming efficiency.

One of the major advantages of the transistor of FIG. **6** is the fact that the transistor is programmed (1) without drawing more than a few microamps of drain current, and (2) without requiring precise regulation of control gate, erase gate or drain voltages. As mentioned above, in the embodiment in which a ramp voltage is applied to the control gate and the embodiment in which $V_T+\Delta$ is applied to the control gate, the control gate voltage has to be precisely controlled to permit programming without drawing too much drain current. In FIG. **6**, programming is achieved without having to tightly regulate the control gate voltage.

Instead of using a resistor as current limiter **36**, in one embodiment, a circuit **51**, comprising a first MOSFET Q3 and a second MOSFET Q4, coupled in a current mirror configuration, provides an MOS equivalent of a resistor (FIG. **7**). A resister R is coupled between VCC and the drain of transistor Q3. The effective resistance

R_{EQ} between the source and drain of transistor Q4 is as follows:

$$R_{EQ} = R_1 \times (W3/L3)/(W4/L4)$$

where W3, L3, W4 and L4 are the channel width of transistor Q3, the channel length of transistor Q3, the channel width of transistor Q4 and the channel length of transistor Q4, respectively, and R1 is the resistance of resistor R. W3, L3, W4 and L4 are selected so that transistor Q4 exhibits a desired amount of resistance. This effective resistance R_{EQ} is typically within the range of 100 KΩ to 2 MΩ, and preferably about 1 MΩ, to permit a drain current less than about 10 μA to flow through floating gate transistor **10**.

In accordance with another embodiment of our invention, drain **14** is coupled to a charge pump **52** while control gate **6** is coupled to a pulse source **54** (FIG. **8**). Pulse source **54** provides a stream of pulses having an amplitude of about 5 volts (VCC), an on-time of 0.1 microseconds and an off-time of 0.9 microseconds. (The waveform provided by pulse source **54** is illustrated in FIG. **9***a*.) Of importance, when control gate **16** is at ground, transistor **10** is off, and drain **14** charges to about 8 volts. When control gate **16** is pulsed, drain **14** is discharged through transistor **10**, and when control gate **16** is again grounded, drain **14** charges back to about 8 volts. FIG. **9***b* illustrates the drain voltage waveform resulting from coupling charge pump **52** to drain **14** and pulsing control gate **16**. The repetitive application of the waveform of FIGS. **9***a* and **9***b* to control gate **16** and drain **14** over about a 1 ms time period is sufficient to program transistor **10**, because at least during time periods P2, the voltage conditions are appropriate for accelerating hot electrons onto floating gate **18**. (If control gate **16** were not pulsed, drain **14** would remain at a low voltage because the charge pump coupled to drain **14** cannot provide a large output current, and the conditions required for hot electron injection would not exist.)

A novel technique for constructing a flash EPROM transistor **101** (FIG. **10***d*) for use with the above described programming technique is described below.

First, a P− silicon substrate **100** is implanted with P type impurities to form a P+layer **102** approximately 0.8 microns thick and having a dopant concentration of between 10^{17} and $10^{18}/cm^3$ (FIG. **10***a*). An insulating layer **104** (typically thermally grown SiO_2) is formed on the wafer, and a heavily doped polysilicon floating gate **105** is formed on insulating layer **104** in a conventional manner. (During formation of floating gate **105**, other floating gates such as floating gates **105′** and **105″** are formed elsewhere on the surface. The description herein only refers to structures within transistor **101**, it being understood that similar structures constituting the rest of a flash EPROM array are formed elsewhere on the wafer.) A photoresist layer **106** is then formed on the wafer and patterned.

Referring to FIG. **10***b*, the wafer is then subjected to an N type ion implantation step to form N+source **108** and drain **110**. One edge **108***a* of source **108** and one edge **110***a* of drain **110** are defined by photoresist **106**, while the other edge **108***b* of source **108** and edge **110***b* of drain **110** are defined by edges **105***a′* and **105***a* of floating gates **105′** and **105**, respectively. This is done for reasons described in U.S. Pat. No. 4,639,893, issued to Boaz Eitan, and incorporated herein by reference.

Photoresist layer **106** is removed, and the wafer is subjected to a diffusion step. The wafer is then subjected to a blanket N type ion implantation step to partially counter-dope a portion **114** of P+layer **102**, so that portion **114** becomes P− material (see FIG. **10***c*). It will be appreciated that at the conclusion of this process step, the transistor channel will include a first area A1 which comprises P− material and a second area A2 which comprises P+material. P+area A2 serves to enhance the transistor programming efficiency, while area A1 is P- material so that the effective threshold voltage of area A1 is about one volt. Of importance, the lateral extent of areas A1 and A2 are self-aligned with the other transistor structures. Thus, it is impossible to misalign the lateral extent of areas A1, A2 and degrade manufacturing yields.

The wafer is then subjected to an oxide etching step (e.g. using HF acid) to remove the exposed portions of insulating layer **104**. An additional insulation layer **116** is then formed on the wafer (e.g. by thermal oxidation). Transistor **101** is completed by forming control gate **120** on the wafer using conventional techniques. (See FIG. **10***d*).

The threshold voltage of area A2 when floating gate **105** is electrically neutral is approximately 3 to 5 volts because of the enhanced channel doping concentration. However, transistor **101** is a flash EPROM. Prior to use, charge is removed from floating gate **105** with an erase gate (not shown in FIGS. **10***a* to **10***d*, but described below) prior to use. This reduces the threshold voltage of area A2 below zero volts. (Although this may result in an inversion region forming under floating gate **105** independently of the voltage at control gate **120**, this will not create a problem since area A1 will only conduct when a high voltage is applied to control gate **120**.) Under these circumstances, transistor **101** stores a zero. Floating gate **105** can then be programmed to raise the threshold voltage of area A2 and to thereby store a one in transistor **101**.

During the process of constructing control gate **120**, an erase gate (not shown in FIGS. **10***a* to **10***d*) is also formed over floating gate **105**, typically outside of the cross section of FIGS. **10***a* to **10***d*. The resulting cell may have a layout as illustrated and described in U.S. Pat. application Ser. No. 07/189,874, entitled "EEPROM WITH IMPROVED ERASE STRUCTURE" filed by Eitan et al. on May 3, 1988, incorporated herein by reference. FIG. **11** illustrates a portion of the layout of an array **200** of flash EPROM cells constructed in accordance with an alternative embodiment of the invention. As can be seen, array **200** includes an array of floating gates **202***a* to **202***h*, source/drain regions **204***a* to **204***c*, control gates **206***a* to **206***d*, tunneling erase gates **208***a*, **208***b*, and field oxide regions **209**. Array **200** is constructed using a staggered virtual grounded architecture. When it is desired to read or program floating gates **202***a* or **202***b*, source/drain region **204***a* serves as a drain while source/drain region **204***b* serves as a source. When it is desired to read or program floating gates **202***c* or **202***d*, source/drain region **204***b* serves as a drain while source/drain region **204***a* serves as a source. Source/drain regions **204***b*, **204***c* similarly serve as a source or a drain to read or program one of floating gates **202***e* to **202***h*. Control gate **206***a* is used to read or program the floating gates within the column comprising floating gates **202***a* and **202***e*. The other control gates are used to read or program the floating gates within other associated columns of floating gates. Erase gate **208***a* is used to erase floating gates **202***a*, **202***c*, **202***e*

and 202*g*, while erase gate 208*b* is used to erase floating gates 202*b*, 202*d*, 202*f* and 202*h*.

In the array of FIG. 11, the floating gates are staggered relative to one another, i.e. the floating gates 202*a* is formed against source/drain region 204*a* while adjacent floating gate 202*c* is formed against source/drain region 202*b*. If floating gates 202*a* and 202*c* were both formed against source/drain region 204*a*, the cell size would have to be increased to permit both floating gates 202*a* and 202*c* to extend underneath erase gate 208*a*. Thus, staggering the floating gates permits the flash EPROM array to be constructed on a small surface area.

An address decoder appropriate for use with the array of FIG. 11 is discussed in U.S. patent application Ser. No. 07/258,926, filed on Oct. 17, 1988 by Syed Ali and incorporated herein by reference. Also see U.S. patent application Ser. No. 07/258,952, filed by Eitan et al. on Oct. 17, 1988.

While the invention has been described with regard to specific embodiments, those skilled in the art will recognize that changes can be made in form and detail without departing from the spirit and scope of the invention. Accordingly, all such changes come within the present invention.

I claim:

1. A method for programming a floating gate transistor, said floating gate transistor comprising a source, a drain spaced apart from said source, said source and drain being of a first conductivity type and formed in a semiconductor region of a second conductivity type, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate, said method comprising the steps of:

applying a programming voltage to said drain and control gate sufficient to cause hot electron injection programming of said transistor; and

ensuring that the programming drain current for said transistor is less than a predetermined value.

2. Method of claim 1 wherein said predetermined value is less than or equal to about 150 μA.

3. Method of claim 1 wherein said predetermined value is less than or equal to about 10 μA.

4. Method of claim 1 wherein said programming drain voltage is provided by a charge pump and said programming drain current is held to a value sufficiently low so that said charge pump can provide said programming drain current.

5. A method for programming a floating gate transistor, said floating gate transistor comprising a source, a drain spaced apart from said source, said source and drain being of a first conductivity type and formed in a semiconductor region of a second conductivity type, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate, said method comprising the steps of:

applying a programming voltage to said drain; and

applying to said control gate a voltage which rises from a first value to a second value such that during the time said voltage at said control gate is rising, electrons are being injected into said floating gate so that the threshold voltage of said transistor increases at a rate which ensures that said transistor does not draw a drain current over a predetermined value during programming.

6. Method of claim 5 wherein the voltage applied to said control gate rises from said first value to said second value over a 0.1 ms time period.

7. Method of claim 1 wherein said step of ensuring comprises the step of applying a voltage to said control gate to keep the drain current below said predetermined value.

8. Method of claim 7 wherein said transistor comprises an erase gate capacitively coupled to said floating gate, said method further comprising the step of raising the voltage at said erase gate.

9. A method for programming a floating gate transistor, said floating gate transistor comprising a source, a drain spaced apart from said source, said source and drain being of a first conductivity type and formed in a semiconductor region of a second conductivity type, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate, said method comprising the steps of:

applying a programming voltage to said drain and control gate;

ensuring that the programming drain current is less than a predetermined value, wherein said step of ensuring comprises the step of providing an electrical resistance between said source and ground, said semiconductor region being grounded.

10. Method of claim 9 wherein said transistor comprises an erase gate capacitively coupled to said floating gate, said method further comprising the step of raising the voltage at said erase gate.

11. Method of claim 10 wherein the voltage at said erase gate is greater than 5 volts during programming.

12. Method of claim 10 wherein the voltage at said erase gate is less than 20 volts during programming.

13. Method of claim 9 wherein said transistor comprises an erase gate which is grounded during programming.

14. A method for programming a floating gate transistor, said transistor comprising a source, a drain, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate, said method comprising the steps of:

coupling a programming voltage generator to said drain; and

repetitively applying pulses to said control gate to thereby cause hot electron injection programming of said transistor.

15. Method of claim 14 wherein said programming voltage generator is a charge pump.

16. Method of claim 14 wherein said transistor comprises an erase gate capacitively coupled to said floating gate, said method further comprising the step of raising the voltage at said erase gate.

17. Method of claim 14 wherein said transistor comprises an erase gate which is grounded during programming.

18. A method for programming a floating gate transistor, said transistor comprising a source, a drain, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a

portion of said floating gate, said method comprising the steps of:

coupling a programming voltage generator to said drain; and

repetitively applying pulses to said control gate,

wherein said programming voltage generator is incapable of generating an output current which said transistor would normally conduct if (1) said floating gate were unprogrammed, (2) said programming drain voltage was applied to said drain, and (3) a programming control gate voltage equal to the amplitude of said pulses was applied to said control gate.

19. Structure comprising:

a floating transistor including a source, a drain spaced apart from said source, said source and drain being of a first conductivity type and formed in a semiconductor region of a second conductivity type, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate; and

means for applying a programming voltage to said drain and control gate to thereby program said transistor by hot electron injection and ensuring that the programming drain current of said transistor is less than a predetermined value.

20. Structure of claim **19** wherein said predetermined value is less than or equal to about 150 μA.

21. Structure comprising:

a floating gate transistor including a source, a drain spaced apart from said source, said source and drain being of a first conductivity type and formed in a semiconductor region of a second conductivity type, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and control gate extending over at least a portion of said floating gate; and

means for applying a programming voltage to said drain and control gate and ensuring that the programming drain current is less than a predetermined value, wherein said means for applying applies to said control gate a voltage which rises from a first value to a second value such that during the time said voltage at said control gate is rising, electrons are being injected into said floating gate so that the threshold voltage of said transistor increases at a rate which ensures that said transistor does not draw a drain current over said predetermined value during programming.

22. Structure of claim **21** wherein said means for applying causes the voltage at said control gate to rise from said first value to said second value over a 0.1 ms time period.

23. Structure of claim **19** wherein an electrical resistance is provided between said source and ground to prevent said programming drain current from exceeding said predetermined value, said semiconductor region being grounded.

24. Structure of claim **23** wherein said transistor comprises an erase gate capacitively coupled to said floating gate, and said means for applying also raises the voltage at said erase gate during programming.

25. Structure comprising:

a floating gate transistor including a source, a drain, a channel extending between said source and drain, a floating gate extending over at least a portion of said channel, and a control gate extending over at least a portion of said floating gate; and

means for repetitively applying pulses to said control gate to thereby program said floating gate transistor by hot electron injection.

26. Structure of claim **25** wherein said transistor includes an erase gate capacitively coupled to said floating gate, said structure further comprising means for raising the voltage at said erase gate during programming.

27. Method of claim **1** wherein said step of ensuring comprises the step of coupling the source of said floating gate transistor to an additional transistor, said additional transistor limiting the current of said floating gate transistor during programming.

28. Method of claim **1** wherein said step of ensuring comprises the step of raising the source voltage of said floating gate transistor during programming.

29. Structure of claim **19** further comprising means for increasing the voltage at said source during programming.

30. Structure of claim **19** further comprising an additional transistor coupled to the source of said floating gate transistor, said additional transistor limiting the programming current of said floating gate transistor.

* * * * *

E.D. Tex. Local Civil Rule 7(d)

(d) Response and Briefing. The response and any briefing shall be contained in one document. A party opposing a motion shall file the response, any briefing and supporting documents within the time period prescribed by Subsection (e) of this rule. A response shall be accompanied by a proposed order conforming to the requirements of Subsection (a) of this rule. Briefing shall contain a concise statement of the reasons in opposition to the motion and a citation of authorities upon which the party relies. In the event a party fails to oppose a motion in the manner prescribed herein, the court will assume that the party has no opposition.

E.D. Tex. Rules of Practice for Patent Cases 3-1

Not later than 10 days before the Initial Case Management Conference with the Court, a party claiming patent infringement must serve on all parties a "Disclosure of Asserted Claims and Infringement Contentions." Separately for each opposing party, the "Disclosure of Asserted Claims and Infringement Contentions" shall contain the following information:

(a) Each claim of each patent in suit that is allegedly infringed by each opposing party;

(b) Separately for each asserted claim, each accused apparatus, product, device, process, method, act, or other instrumentality ("Accused Instrumentality") of each opposing party of which the party is aware. This identification shall be as specific as possible. Each product, device, and apparatus must be identified by name or model number, if known. Each method or process must be identified by name, if known, or by any product, device, or apparatus which, when used, allegedly results in the practice of the claimed method or process;

(c) A chart identifying specifically where each element of each asserted claim is found within each Accused Instrumentality, including for each element that such party contends is governed by 35 U.S.C. § 112(6), the identity of the structure(s), act(s), or material(s) in the Accused Instrumentality that performs the claimed function;

(d) Whether each element of each asserted claim is claimed to be literally present or present under the doctrine of equivalents in the Accused Instrumentality;

(e) For any patent that claims priority to an earlier application, the priority date to which each asserted claim allegedly is entitled; and

(f) If a party claiming patent infringement wishes to preserve the right to rely, for any purpose, on the assertion that its own apparatus, product, device, process, method, act, or other instrumentality practices the claimed invention, the party must identify, separately for each asserted claim, each such apparatus, product, device, process, method, act, or other instrumentality that incorporates or reflects that particular claim.

E.D. Tex. Rule of Practice for Patent Cases 3-6

(a) Leave not required. Each party's "Infringement Contentions" and "Invalidity Contentions" shall be deemed to be that party's final contentions, except as set forth below.

> (1) If a party claiming patent infringement believes in good faith that the Court's Claim Construction Ruling so requires, not later than 30 days after service by the Court of its Claim Construction Ruling, that party may serve "Amended Infringement Contentions" without leave of court that amend its "Infringement Contentions" with respect to the information required by Patent R. 3-1(c) and (d).
>
> (2) Not later than 50 days after service by the Court of its Claim Construction Ruling, each party opposing a claim of patent infringement may serve "Amended Invalidity Contentions" without leave of court that amend its "Invalidity Contentions" with respect to the information required by P. R. 3-3 if:
>
> > (A) a party claiming patent infringement has served "Infringement Contentions" pursuant to P. R. 3-6(a), or
> >
> > (B) the party opposing a claim of patent infringement believes in good faith that the Court's Claim Construction Ruling so requires.

(b) Leave required. Amendment or supplementation any Infringement Contentions or Invalidity Contentions, other than as expressly permitted in P. R. 3-6(a), may be made only by order of the Court, which shall be entered only upon a showing of good cause.

Fed. R. Civ. P. Rule 16(f)

(f) SANCTIONS.

(1) *In General.* On motion or on its own, the court may issue any just orders, including those authorized by Rule 37(b)(2)(A)(ii)–(vii), if a party or its attorney:

(A) fails to appear at a scheduling or other pretrial conference;

(B) is substantially unprepared to participate—or does not participate in good faith—in the conference; or

(C) fails to obey a scheduling or other pretrial order.

(2) *Imposing Fees and Costs.* Instead of or in addition to any other sanction, the court must order the party, its attorney, or both to pay the reasonable expenses—including attorney's fees—incurred because of any noncompliance with this rule, unless the noncompliance was substantially justified or other circumstances make an award of expenses unjust.

Fed. R. Civ. P. Rule 37(b)

(b) FAILURE TO COMPLY WITH A COURT ORDER.

(1) *Sanctions Sought in the District Where the Deposition Is Taken.* If the court where the discovery is taken orders a deponent to be sworn or to answer a question and the deponent fails to obey, the failure may be treated as contempt of court. If a deposition-related motion is transferred to the court where the action is pending, and that court orders a deponent to be sworn or to answer a question and the deponent fails to obey, the failure may be treated as contempt of either the court where the discovery is taken or the court where the action is pending.

(2) *Sanctions Sought in the District Where the Action Is Pending.*

(A) *For Not Obeying a Discovery Order.* If a party or a party's officer, director, or managing agent—or a witness designated under Rule 30(b)(6) or 31(a)(4)—fails to obey an order to provide or permit discovery, including an order under Rule 26(f), 35, or 37(a), the court where the action is pending may issue further just orders. They may include the following:

(i) directing that the matters embraced in the order or other designated facts be taken as established for purposes of the action, as the prevailing party claims;

(ii) prohibiting the disobedient party from supporting or opposing designated claims or defenses, or from introducing designated matters in evidence; (iii) striking pleadings in whole or in part;

(iv) staying further proceedings until the order is obeyed;

(v) dismissing the action or proceeding in whole or in part;

(vi) rendering a default judgment against the disobedient party; or

(vii) treating as contempt of court the failure to obey any order except an order to submit to a physical or mental examination.

(B) *For Not Producing a Person for Examination.* If a party fails to comply with an order under Rule 35(a) requiring it to produce another person for examination, the court may issue any of the orders listed in Rule 37(b)(2)(A)(i)–(vi), unless the disobedient party shows that it cannot produce the other person.

(C) *Payment of Expenses.* Instead of or in addition to the orders above, the court must order the disobedient party, the attorney advising that party, or both to pay the reasonable expenses, including attorney's fees, caused by the failure, unless the failure was substantially justified or other circumstances make an award of expenses unjust.

United States Court of Appeals
for the Federal Circuit

Keranos, LLC v. Analog Devices, Inc., 2014-1360, -1500

CERTIFICATE OF SERVICE

I, Michelle G. Breit, being duly sworn according to law and being over the age of 18, upon my oath depose and say that:

On **August 11, 2014**, I have authorized Counsel Press to electronically file the foregoing **Brief for Plaintiff-Appellant (confidential and non-confidential versions)** with the Clerk of Court using the CM/ECF System, which will serve via e-mail notice of such filing to all counsel registered as CM/ECF users, including any of the following:

Jeffrey D. Mills
Brian C. Banner
Adam Conrad
Bruce W. Slayden, II
King & Spalding LLP
401 Congress Avenue, Suite 3200
Austin, TX 78701
512-457-2027
jmills@kslaw.com
bbanner@kslaw.com
aconrad@kslaw.com
bslayden@kslaw.com
Counsel for Analog Devices, Inc., IBM Corporation, Intel Corporation, National Semiconductor Corporation, NXP Semiconductors USA, Inc., and Texas Instruments, Inc.,

John R. Lane
Fish & Richardson, P.C.
One Houston Center, Suite 2800
1221 McKinney
Houston, TX 77010
713-654-5300
jlane@fr.com
Counsel for Apple Inc.

By agreement between the parties, confidential copies (partly redacted for each Appellee in compliance with the protective order) will be emailed to the above counsel on this date.

Upon acceptance by the Court of the e-filed document I will cause six confidential paper copies will be filed with the Court, within the time provided in the Court's rules.

Dated: August 11, 2014

/s/ Michelle G. Breit
Michelle G. Breit
Attorney for Plaintiff – Appellant
Keranos, LLC

CERTIFICATE OF COMPLIANCE WITH TYPE-VOLUME LIMITATION, TYPEFACE REQUIREMENTS AND TYPE STYLE REQUIREMENTS

1. This brief complies with the type-volume limitation of Federal Rule of Appellate Procedure 32(a)(7)(B).

__X__ The brief contains 13,827 words, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(a)(7)(B)(iii),or

____ The brief uses a monospaced typeface and contains _____ lines of text, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(a)(7)(B)(iii).

2. This brief complies with the typeface requirements of Federal Rule of Appellate Procedure 32(a)(5) and the type style requirements of Federal Rule of Appellate Procedure 32(a)(6).

__X__ The brief has been prepared in a proportionally spaced typeface using MS Word 2013 in a ___ point _____ font or

____ The brief has been prepared in a monospaced typeface using _____ _____in a ___ characters per inch_____ font.

Dated: August 11, 2014

/s/ Michelle G. Breit
Michelle G. Breit
Attorney for Plaintiff – Appellant
Keranos, LLC